UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author Li, Cheryl Suet-Yee

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOAN

Theses may not be lent to individuals, but the University Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: The Theses Section, University of London Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the University of London Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

A. Before 1962. Permission granted only upon the prior written consent of the author. (The University Library will provide addresses where possible).

B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.

C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.

D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

☐ This copy has been deposited in the Library of UCL

☐ This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.

Viral and Host Gene Expression Patterns in Human Herpesvirus Type-7 Infection of T-Cells

Cheryl Suet-Yee Li

Submitted to the University of London for the degree of Doctor of Philosophy

December 2005

Viral Genomics and Bioinformatics Group
Department of Immunology and Molecular Pathology
Windeyer Institute of Medical Sciences
University College London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Human herpesvirus type 7 (HHV-7), with its high prevalence and apparently non-pathogenic nature, is believed to be a virus very well adapted to humans. It is therefore useful to study the transcription patterns of the virus and the host to understand the life cycle and pathogenesis (or the lack of it) of HHV-7 infection. In order to obtain the global pattern of host and viral gene expression, we have made a DNA microarray containing probes that target all 86 predicted HHV-7 open reading frames and approximately 1000 cellular genes. This is the first description of an integrated host-pathogen microarray for HHV-7.

The microarrays were used to follow the gene expression patterns over a 72-hour time course infection of a $CD4^+$ T-cell line (Sup-T1) by HHV-7 *in vitro*, which revealed an overall increase of HHV-7 transcripts over time, as well as the coordinated transcription profiles for groups of viral genes with similar functions. HHV-7 genes predicted to be immediate-early transcription factors were the first genes to be expressed, followed by genes involved in viral DNA replication, and at later stages, the genes encoding structural proteins and proteins of packaging functions were expressed. Temporal classification was assigned for 68 HHV-7 ORFs, of which a selection was confirmed by RT-PCR in the presence of metabolic inhibitors of *de novo* protein synthesis and viral DNA replication. Both gene expression profiling and inhibitor treatment demonstrated that the lytic gene expression of HHV-7 was temporally regulated, consistent with the classification of viral gene expression into immediate-early, early and late transcripts in other herpesviruses.

ACKNOWLEDGEMENTS

This is probably my favourite page in this thesis, when I finally have the chance to breathe and to reflect on the past few years. Throughout this PhD, there have been difficult and wonderful times, sometimes simultaneously, and it has been a remarkable experience indeed. I would like to thank many people who have provided help towards the production of this thesis, and who have shown me kindness thoughout the process.

First of all, I am truly grateful for having Paul Kellam as my supervisor. Thank you Paul, for your guidance, patience, encouragement, and for sharing your vision and expertise in genomic science with all of us in VGB. I thank Mary Collins, Robin Weiss, Aíne McKnight, Ari Fassati and Vince Emery for their invaluable suggestions and inspirations. The HHV-7 virus work could not have happened without the help of Duncan Clark from the Royal Free Virology Department, and the production of the microarrays would be impossible without the contributions from Ed Tsao and Jane Rasaijaah. I also thank Nicola Gilbert, Liz Thomson, Carolyn McElvaney and John Cheshire for their generous help all these years. This PhD would be challenging but no fun without all my colleagues in the VGB group, who provided a vibrant environment for research. I am thankful to the MRC for funding this PhD project.

Finally, thanks to my parents for their understanding and patience to a daughter who never calls, visits only once a year, and for their support in my choice of career. Last but not least, thanks to Jacob, for sharing everything with me.

Thanks be to God. May His name be praised.

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

°C	Degree Celsius
2d-gel	2-dimentional gel
Acc. No.	Accession number
ANOVA	Analysis of variance
aRNA	Amplified RNA
B-cell	Bone marrow-derived lymphocyte
BCG	Bacillus of Calmette and Guérin
BHV	Bovine herpesvirus
BLAST	Basic local alignment search tool
BLASTN	BLAST nucleotide
BSA	Bovine serum albumin
C terminus	Carboxyl (COOH)-terminus
CBMC	Cord blood mononuclear cell
CCV	Channel catfish virus
cDNA	Complementary DNA
CFS	Chronic fatigue syndrome
CGH	Comparative genome hybridisation
CHX	Cycloheximide
cm^2	Squared centimetre
CMV	Cytomegalovirus
CO_2	Carbon dioxide
CPE	Cytopathic effect
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
ddATP	Dideoxyadenosine triphosphate
ddCTP	Dideoxycytidine triphosphate
ddGTP	Dideoxyguanosine triphosphate
ddUTP	Dideoxyuridine triphosphate
dGTP	Deoxyguanosine triphosphate
dITP	Deoxyinosine triphosphate
DC	HHV-7 strain DC
DMEM	Dulbeccos modified Eagle's media
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
dpi	Day post infection
DR	Direct repeat
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
dUTPase	Deoxyuridine triphosphatase
E	Early
E. coli	Escherichia coli
EBNA	Epstein-Barr nuclear antigen
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetraacetic acid
EHV	Equine herpesvirus
FACS	Fluorescence activated cell sorting
FBS	Foetal bovine serum
FCS	Foetal calf serum
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
FSC	Forward scatter
g	Gram
gB	Glycoprotein B
gH	Glycoprotein H
gJ	Glycoprotein J
gL	Glycoprotein L

gM	Glycoprotein M
G-protein	Guanine nucleotide binding proteins
h	Hour
HCL	Hierarchical clustering
HCMV	Human cytomegalovirus
HHV	Human herpesvirus
HIV	Human immunodeficiency virus
hpi	Hour post infection
HPV	Human papilloma virus
HSV	Herpes simplex virus
ICP	Infected cell polypeptides
IE	Immediate early
IFA	Immunofluorescence assay
IgG	Immunoglobulin G
IPTG	Isopropylthio--D-galactoside
IUC	Infectious units count
IVT	*In vitro* transcription
JI	HHV-7 strain JI
kb	Kilobase
KSHV	Kaposi's sarcoma-associated herpesvirus
L	Late
LB	Luria-Bertani
log_2	Log base 2
LOWESS	Local weighted scatterplot smoothing
M. tuberculosis	Mycobacterium tuberculosis
M/A	Scatter plot of Cy5/Cy3 ratio versus geometric mean
mA	MilliAmpere
MCMV	Murine cytomegalovirus
mCP	Minor capsid protein
MCP	Major capsid protein
MeV	Multiexperiment viewer
µg	Microgram
MHV	Murine gammaherpesvirus
MIAME	Minimal information about a microarray experiment
MIDAS	Microarray Data Analysis System
MIE	Major immediate-early
min	Minute
µl	Microlitre
ml	Millilitre
µM	Micromolar
mM	Millimolar
µmol	Micromole
MOI	Multiplicity of infection
mRNA	Messenger RNA
ng	Nanogram
nm	Nanometre
N-terminal	Amino (NH_2)-terminal
OBP	Origin-binding protein
OD	Optical density
ORF	Open reading frame
oriLyt	Origin of lytic replication
P/S	penicillin and streptomycin
PAA	Phosphonoacetic acid
PBL	Peripheral blood lymphocyte
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Phosphonoformic acid
pmol	Picomole
PMT	Photomultiplier tube
PR	Pityriasis rosea
PTFE	Polytetrafluoroethylene

RK	HHV-7 strain RK
RNA	Ribonucleic acid
RNase	Ribonuclease
RPMI	Roswell Park Memorial Institute
RT-PCR	Reverse transcription-PCR
s	Second
SAM	Significance analysis of microarrays
SARS-CoV	Severe acute respiratory syndrome-associated coronavirus
SDS	Sodium dodecyl sulphate
SNP	Single nucleotide polymorphism
SNR	Signal-to-noise ratio
SOM	Self-organising map
SSC	Side scatter
SSC	Saline sodium citrate
SSPE	Saline-sodium phosphate-EDTA
TAE	Tris-acetate-EDTA
T-cell	Thymus-derived lymphocyte
$TCID_{50}$	Tissue culture infectious $dose_{50}$
TE	Tris-EDTA
TIGR	The Institute for Genome Research
TNF	Tumour necrosis factor
TP	Transport protein
TRAIL	TNF-related apoptosis-inducing ligand
TRAIL-R	TRAIL receptor
T-RNA	Total RNA
U	Unique region
UK	United Kingdom
UL	Unique long region
US	Unique short region
UV	Ultra-violet
V	Ig variable region
V	Volt
v/v	Volume by volume
v-fos	Viral homologue of the fos family of transcription factors
vIRFs	Viral interferon regulatory factors
VZV	Varicella zoster virus
w/v	Weight by volume
X-gal	5-bromo-4-chloro-3-indolyl--D-galactoside

CHAPTER 1

INTRODUCTION

1.1 Human Herpesvirus 7

1.1.1 A brief history of human herpesvirus type-7

Human herpesvirus type-7 (HHV-7) was discovered in 1990, when Frenkel and colleagues observed abnormal cytopathic effect (CPE) in primary T-cells, obtained from the peripheral blood of a healthy individual, RK, under T-cell activating conditions. A previously unidentified virus was isolated from the culture and typical herpesvirus virions were revealed by electron microscopy. The virion had a diameter of 170 nm, consisting an electron dense cylindrical core, capsid, tegument, and envelope (Frenkel et al., 1990; Secchiero et al., 1994). Restriction pattern and DNA hybridisation analysis showed the virus to be distinct from the six known human herpesviruses at the time (Frenkel et al., 1990), and was therefore designated the seventh human herpesvirus. HHV-7 was classified as a member of the Roseolovirus genus of the betaherpesvirus subfamily. Soon after, laboratory cultivation of HHV-7 in a continuous cell line (SupT1) was reported (Berneman et al., 1992a), enabling easier production of the virus and advancing research, including the identification of a cellular receptor (Lusso et al., 1994) and deciphering the complete genome sequence (Megaw et al., 1998; Nicholas, 1996).

1.1.2 Human herpesvirus

To date, 8 human herpesviruses have been identified. They share some common features, for example they are large viruses of 120 - 300 nm in diameter, with a complex virion structure composed of an envelope, a tegument and an icosahedral capsid that encase a relatively long (120 – 250 kb), linear, double stranded DNA genome. The herpesvirus genomes consist of 7 blocks of core genes, which are preserved in all members albeit arranged in different order (Roizman, 2001b). All herpesviruses encode a large number of enzymes involved in nucleic acid metabolism, DNA synthesis, and protein processing. The expression of these genes in productive infection results in progeny virus release and the death of host cell. All herpesviruses can establish a latent infection in the host. Viral genomes exist as closed circular molecules in the host cell

during latency, and viral gene expression is minimised (Roizman, 2001b). Nonetheless, each herpesvirus has unique properties and some viruses are more similar than others. Herpesviruses are classified into three subfamilies, alphaherpesvirus, betaherpesvirus and gammaherpesvirus, based on biological properties (Table 1.1). The closest relative of HHV-7 is HHV-6, discovered in 1986 (Salahuddin et al., 1986), with amino acid sequence similarities ranging from 41 to 75% between products of the core gene homologues (Nicholas, 1996). These two viruses belong to the genus of Roseolovirinae.

Table 1.1 Human herpesviruses - classification and disease associations

Common name of virus and subfamily	Human herpesvirus nomenclature	Disease association
Alphaherpesviruses		
Herpes simplex virus type 1 (HSV-1)	HHV-1	Oropharygeal herpes (cold sore)
Herpes simplex virus type 2 (HSV-2)	HHV-2	Genital herpes
Varicella zoster virus (VZV)	HHV-3	Varicella (chicken pox) Herpes zoster
Betaherpesviruses		
Human cytomegalovirus (HCMV)	HHV-5	Mononucleosis, Retinitis Complications in immunocompromised/transplant
Human herpesvirus 6A	HHV-6A	Exanthem subitum (rare)
Human herpesvirus 6B	HHV-6B	Exanthem subitum Transplant complications
Human herpesvirus 7	HHV-7	Exanthem subitum
Gammaherpesviruses		
Epstein-Barr virus (EBV)	HHV-4	B-cell lymphoma Nasopharyngeal carcinoma
Kaposi's sarcoma-associated virus (KSHV)	HHV-8	Kaposi's sarcoma (KS) Primary effusion lymphoma Multicentric Castleman's disease

1.1.3 Morphology and culture *in vitro*

HHV-7 has been cultured *in vitro* in the $CD4^+$ immature T-cell line SupT1 (Berneman et al., 1992a), activated $CD4^+$ T-cells purified from peripheral blood lymphocytes (Secchiero et al., 1994), peripheral blood mononuclear cells (PBMCs) (Berneman et al., 1992b) and cord blood mononuclear cells (CBMCs) (Ablashi et al., 1998). HHV-7 infection has a 4-day cycle in cord blood lymphocytes (Black et al., 1997). Infection in SupT1 is best maintained in medium supplemented with 5% foetal calf serum (FCS) (Black et al., 1997), probably because a lower FCS concentration (optimum for cells 10%) facilitates virus propagation by retarding cell division. Light microscopy of infected SupT1 and primary T-cells reveal refractile, dense, vacuolated cells and giant multinucleated syncytia (Ablashi et al., 1998; Secchiero et al., 1997a), indicative of viral induced cytopathological effect (CPE). HHV-7 nucleocapsids are observed in the nucleus and cytoplasm of infected cells (Ablashi et al., 1998), and mature virions are seen in extracellular space (Secchiero et al., 1994; Secchiero et al., 1997a). Infected SupT1 also elongate and loosely attach to the culture flask. These flask-attached SupT1 cells contain HHV-7 nucleocapsids and are capable of passing on infection (Ablashi et al., 1998). During the course of *in vitro* infection, viral load increases with time while the viable cell count decreases (Secchiero et al., 1997a). The infected, multinucleated cells die by necrosis, while apparently uninfected, small cells in the same culture undergo apoptosis (Secchiero et al., 1998b; Secchiero et al., 1997a). Both apoptosis and necrosis of these cells were inhibited by PFA (phosphonoformic acid), a viral DNA synthesis inhibitor. Ablashi *et al* suggested HHV-7 is a cell-associated virus hence only released by cell lysis such as freeze-thawing (Ablashi et al., 1998), whereas Secchiero *et al* reported extracellular virus particles by electron microscopy (Secchiero et al., 1994). Various strains of HHV-7 hav been described in the literature, summarised in Table 1.2.

Table 1.2 Strains of HHV-7

Strain	Genome	Source	Titre ($TCID_{50}$)	Particles/ml	Reference [b]
AL		PBMCs of healthy adult			1
CZ		Infant with febrile syndrome			2
H7-2		PBMC/saliva of children	10^3	1.8×10^9	3
H7-3		PBMC/saliva of children	10^3	1.7×10^9	3
H7-4		PBMC/saliva of children	10^4-10^5	3.8×10^9	3
JB	Same restriction pattern to JI	PBMC of adult with CFS[a]	10^3	1.5×10^9	1, 3
JHC		CBMC of adult	10^3	3.9×10^9	3
JI	144861 bp	PBMC of adult with CFS[a]			1, 4, 5
KHR		PBMC/saliva of children (Japan)	10^4-10^5	1.9×10^9	3
RK	153080 bp	PBMC of healthy adult			6, 7

[a] CFS, chronic fatigue syndrome

[b] 1, Secchiero *et al.*, 1994; 2, Menegazzi *et al.*, 1999; 3, Ablashi *et al.*, 1998; 4, Berneman *et al.*, 1992b; 5, Nicholas, 1996; 6, Frenkel *et al.*, 1990, 7, Megaw *et al.*, 1998

1.1.4 Genome

Two strains of HHV-7, JI (Nicholas, 1996) and RK (Megaw et al., 1998) have been fully sequenced and reveal double stranded DNA genomes of approximately 145 and 153 kb, respectively. The HHV-7 genome consists of a central unique region (U) flanked by terminal direct repeats (DR) on either end (Figure 1.1). The genome contains reiterated sequences in the DR regions (T1 and T2) and in the unique region (R1 and R2). These vary in length and copy number between strains, hence account for the difference in overall genome length. Otherwise the genomes of the two strains are highly conserved, differing only by 0.1% (179 bp). Genome analyses have identified the presence of conserved genes, arranged in 7 core gene blocks in the unique region (Figure 1.1), a feature conserved among herpesviruses (Figure 1.2). These core genes encode products involved in regulatory functions, DNA replication and repair, nucleotide metabolism and virion assembly. Amino acid sequences of the core genes have the highest identity (41-75%) to HHV-6, and they are in the same order and orientation as HCMV and HHV-6. The genome of HHV-7 is predicted to have at least 86 open reading frames (ORFs), some with predicted functions, by sequence homology to HHV-6 (Gompels et al., 1995) and HCMV (Chee et al., 1990) (Table 1.3). An extra 14 ORFs has been suggested (Nicholas, 1996) but later considered to be either introns of adjacent ORFs or unlikely to encode protein (Megaw et al., 1998). Of the 86 ORFs there are 84 different genes; 2 ORFs in the 3'-end DR region (DR1' and DR6') are identical to their counterparts in the 5'-end (DR1 and DR6). Up to 11 genes in HHV-7 may be spliced (Megaw et al., 1998). Two regions in the genome (U17-U19 and U86-U90) display conserved layout and splicing pattern to the two IE loci in HCMV (Pellett, 2001), although the amino acid sequences are diverged. 8 ORFs in the HHV-7 genome (DR6, U2, U3, U7, U8, U17, U25, U95) are members of the HCMV US22 gene family of duplicated and diverged genes. Except for 7 genes specific to the roseoloviruses (U20, U21, U23, U24, U26, U100), the rest of the HHV-7 ORFs all have counterparts in HCMV. Amongst the roseoloviruses, HHV-7 encodes U53.5 that has no homologue in HHV-6, and a duplicated gene (U55A and U55B) that is present in HHV-6 as a single gene (U55); otherwise there are no known genes unique to HHV-7 and absent in HHV-6. In contrast, HHV-6 encodes at least 5 genes (DR3, U6, U22, U83, U94) that have no homologues in HHV-7. HHV-7 therefore has a smaller genome and encodes fewer

genes than HHV-6 and HCMV. The origin of lytic replication (*oriLyt*) is identified between U41 and U42 (Figure 1.1).

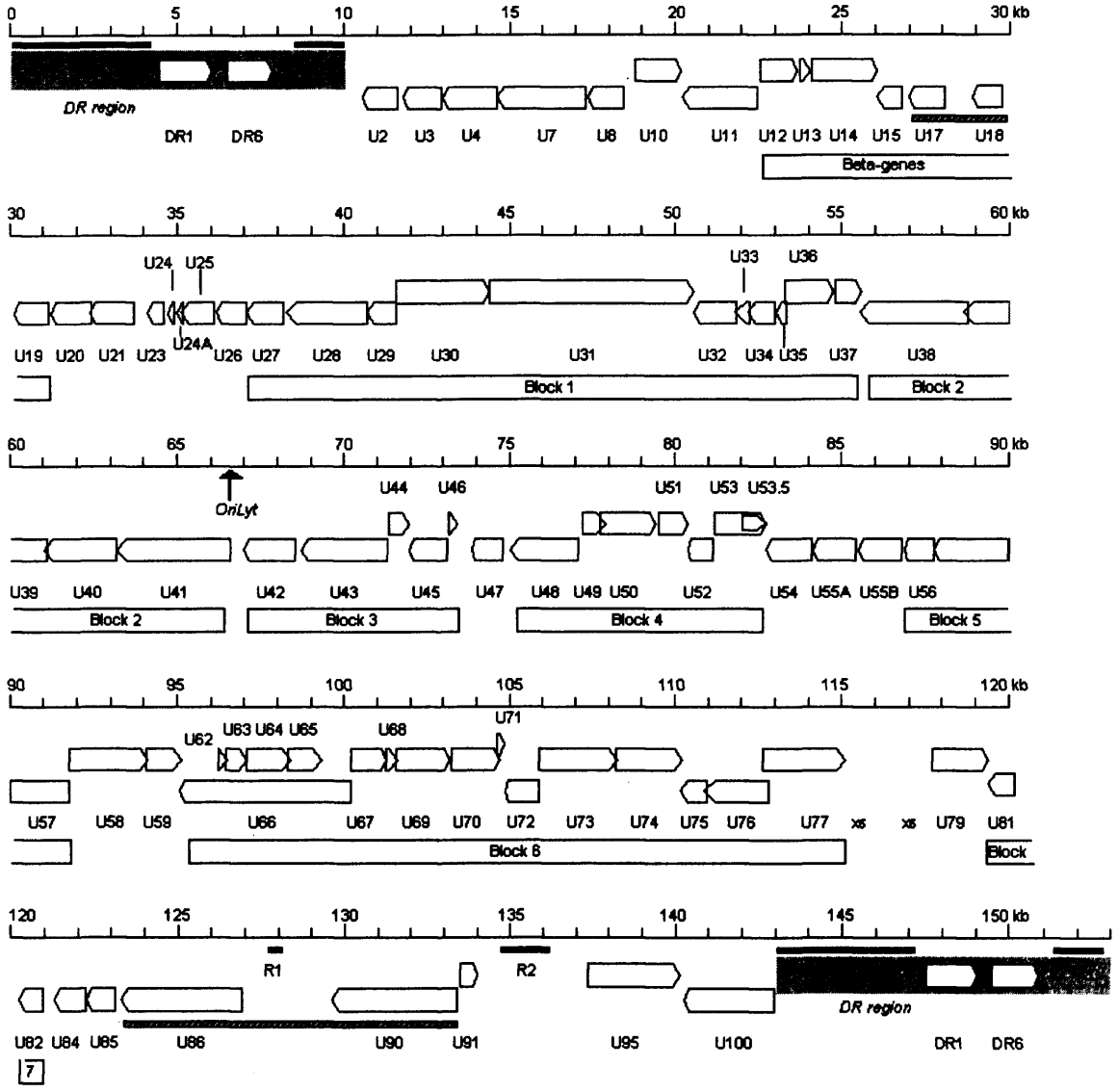


Figure 1.1 HHV-7 (RK) genome arrangement. Shaded region denotes the terminal direct repeats. Reiterated sequences (T1, T2, R1, R2) are labelled and shown as black rectangles. Predicted open reading frames (prefixed by U, DR) are shown as block arrows in the direction of expression, above their gene names. The origin of lytic replication (oriLyt) is indicated by black arrow. Scale in kilobases. The two regions homologous to HCMV immediately early loci are shown as striped rectangles. The 7 herpesvirus core gene blocks and 1 block conserved in betaherpesviruses are labelled and shown as white rectangles.

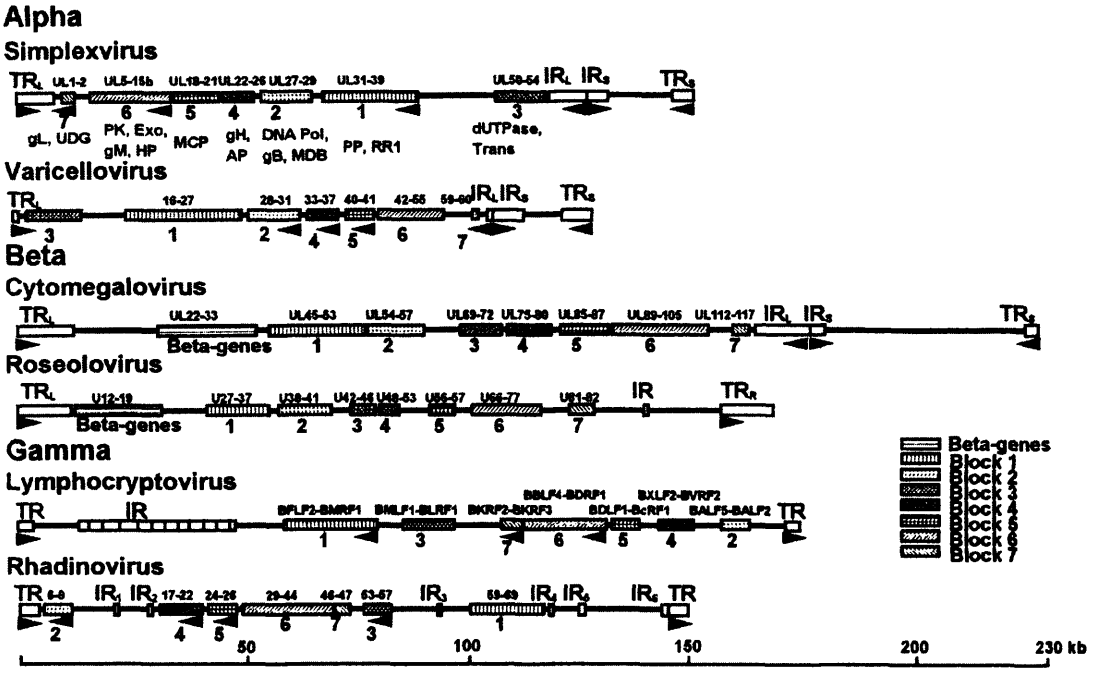


Figure 1.2 Conserved gene blocks of the Alpha-, Beta-, and Gammaherpesvirinae. Human herpesvirus-1 (HHV-1) and HHV-2 belong to the genus Simplexvirus; HHV-3, Varicellovirus; HHV-4, Lymphocryptovirus; HHV-5, Cytomegalovirus; HHV-6 and HHV-7, Roseolovirus; HHV-8, Rhadinovirus. The seven conserved herpesvirus sequence blocks (block 1 through block 7), plus an additional sequence block (beta-genes) found only in betaherpesviruses, are diagramed. ORFs bounding the conserved sequence blocks are indicated. Black arrows, gene blocks inverted relative to the betaherpesviruses; TRL, IRL, IRS, and TRS, inverted repeats flanking the L or S component, respectively; Hatched arrows, directionality of terminal repeats. Gene product abbreviations: gL, glycoprotein L; UDG, uracil-DNA glycosylase; PK, phosphotransferase; Exo, alkaline exonuclease; gM, glycoprotein M; HP, helicase/primase complex; MCP, major capsid protein; gH, glycoprotein H; AP, protease; DNA Pol, DNA polymerase; gB, glycoprotein B; PP, polymerase-associated processivity

Table 1.3 Predicted functions of HHV-7 ORFs and betaherpesvirus homologues

HHV-7 ORF (strain [a])		Homologues [b]		Predicted function	Ref [c]
(JI)	(RK)	HHV-6	HCMV		
H1				Unlikely to encode protein	2
DR1	DR1 ex1	DR1		US22 gene family	1,2
DR2	DR1 ex2	DR2		US22 gene family	1,2
H2				Unlikely to encode protein	2
DR6	DR6 ex1	DR6		US22 gene family	1,2
DR7	DR6 ex2	DR7		US22 gene family	1,2
H3				Unlikely to encode protein	2
H4				Unlikely to encode protein	2
U2	U2	U2	UL23	US22 gene family	1,2
U3	U3	U3	UL24	US22 gene family	1,2
U4	U4	U4	UL27	Related to U7 exon 2	2
U5/7	U7 ex2	U5	UL27	US22 gene family Related to U4	1,2 2
	U7 ex1	U7	UL28	US22 gene family	1,2
U8	U8	U8	UL29	US22 gene family	1,2
U10	U10	U10	UL31	Nuclear phosphoprotein	3
U11	U11	U11	UL32	Structural phosphoprotein	1,2
U12	U12	U12	UL33	G-protein coupled receptor Chemokine receptor	1,2 1
U13	U13	U13		-	1,2
U14	U14	U14	UL35	HCMV UL25/35 gene family, antigenic HHV-7 pp85 Tegument phosphoprotein	4 3
U15	U15	U15		HCMV UL25/35 gene family	1,2
U16	U17 ex2	U16	UL36x2	IE-B, transactivator US22 gene family	1,2 2
U17Ex	U17 ex1	U17Ex	UL36x1	IE-B, transactivator US22 gene family	1,2 2
U17A		U17A		Unlikely to encode protein	2
U18				IE-B, homologue to HCMV IE protein	1
U18	U18	U18	UL37x3	IE-B membrane glycoprotein	2
U19	U19	U19	UL38	IE-B protein	1,2
U20	U20	U20		Probable membrane glycoprotein Ig gene family?	1,2 1
U21	U21	U21		Probable membrane glycoprotein	1,2
U23	U23	U23		Probable membrane glycoprotein EHV-1 gJ homologue	1,2 1
U24	U24	U24		Glycoprotein Contains hydrophobic domain	1 2
U24A	U24A	U24A		Contains hydrophobic domain	2
U25	U25	U25	UL43	US22 gene family Transactivator	1,2 1
U26	U26	U26		-	1,2
U27	U27	U27	UL44	DNA polymerase processivity factor	1,2
U28	U28	U28	UL45	Ribonucleotide reductase (large subunit)	1,2
U29	U29	U29	UL46	Minor capsid protein (mCP) Component of intercapsomeric triplex	1,2 2
U30				Capsid assembly, myosin	1
U30	U30	U30	UL47	Tegument protein	2

HHV-7 ORF (strain [a])		Homologues [b]		Predicted function	Ref [c]
(JI)	(RK)	HHV-6	HCMV		
U31	U31	U31	UL48	Large tegument protein	1,2
U32	U32	U32	UL48.5	Capsid protein (located on tips of hexons)	2
U33	U33	U33	UL49	Virion protein	1,2
U34	U34	U34	UL50	Probable virion protein Membrane-associated phosphoprotein	1 2
U35	U35	U35	UL51	Role in DNA packaging	2
U36	U36	U36	UL52	Probable virion protein Role in DNA packaging	1 2
U37	U37	U37	UL53	Putative phosphoprotein	4
U38	U38	U38	UL54	DNA polymerase Catalytic subunit of DNA polymerase	1 2
U39	U39	U39	UL55	Envelope glycoprotein B (gB)	1,2
U40	U40	U40	UL56	Transport protein TP Role in DNA packaging	1 2
U41	U41	U41	UL57	Major DNA-binding protein Single stranded DNA-binding protein	1,2 2
U42	U42	U42	UL69	Transactivator Post-translational regulator of gene expresson	1 2
U43	U43	U43	UL70	Component of DNA helicase-primase complex, primase	1,2
U44	U44	U44	UL71	-	1,2
U45	U45	U45	UL72	dUTPase	1
U46	U46	U46	UL73	Membrane protein	2
U47	U47	U47	UL74	Envelope glycoprotein O (Imanishi et al.)	1,2
U48	U48	U48	UL75	Envelope glycoprotein H (gH)	1,2
U49	U49	U49	UL76	Fusion protein	1
U50	U50	U50	UL77	Virion protein Role in DNA packaging	1 2
U51	U51	U51	UL78	G-protein coupled receptor OpioidR homologue	1,2 1
U52	U52	U52	UL79	-	1,2
U53	U53	U53	UL80	Protease/assembly protein N-terminal protease domain	1,2 2
	U53.5			Major capsid scaffold protein	2
U54	U54	U54	UL82/83	Tegument protein transactivator Virion transactivator	1 2
U55A	U55A	U55	UL84	Probable replication function Related to U55B	1 2
U55B	U55B	U55	UL84	Probable replication function Related to U55A	1 2
U56	U56	U56	UL85	Capsid protein Component of intercapsomeric triplex	1,2 2
U57	U57	U57	UL86	Major capsid protein (MCP) Forms hexons and pentons	1,2 2
U58	U58	U58	UL87	-	1,2
U59	U59	U59	UL88	-	1,2
U60	U66 ex2	U60	UL89x2	Late spliced gene (U60/U66) DNA packaging Putative terminase	1,2 2
U62	U62	U62	UL91	-	1,2
U63	U63	U63	UL92	-	1,2
U64	U64	U64	UL93	Role in DNA packaging, tegument protein	2

HHV-7 ORF (strain [a])		Homologues [b]		Predicted function	Ref [c]
(JI)	(RK)	HHV-6	HCMV		
U65	U65	U65	UL94	Tegument protein	2
U66	U66 ex1	U66	UL89x1	Late spliced gene (U60/U66) DNA packaging Putative terminase	1,2 2
U67	U67	U67	UL95	-	1,2
U68	U68	U68	UL96	-	1,2
U69 U69	U69	U69	UL97	Phosphotransferase Serine-threonine protein kinase, tegument protein	1 2
U70	U70	U70	UL98	Alkaline exonuclease Deoxyribonuclease, maturation/packaging of DNA	1 2
U71	U71	U71		Myristylated tegument protein	2
U72	U72	U72	UL100	Integral membrane protein (gM) Role in virion envelopment	1,2 2
U73	U73	U73		Origin-binding protein (OBP) Helicase	1,2 2
U74	U74	U74	UL102	Component of DNA helicase/primase complex	1,2
U75	U75	U75	UL103	-	1,2
U76	U76	U76	UL104	Probable virion protein Minor capsid protein, DNA packaging	1 2
U77	U77	U77	UL105	Component of DNA helicase/primase complex Helicase	1,2 2
H5				Unlikely to encode protein	2
U79	U79 ex1	U79	UL112	HCMV replicaton, spliced (UL112/UL113) Probable role in DNA replication	1 2
H6	U79 ex2			HHV-6 U79 homology (C terminus) Probable role in DNA replication	1 2
U80	U79 ex3	U80	UL113	HCMV replicaton, spliced (UL112/UL113) Probable role in DNA replication	1 2
U81	U81	U81	UL114	Uracil-DNA glycosylase	1,2
U82	U82	U82	UL115	Envelope glycoprotein L (gL)	1,2
U84	U84	U84	UL117	Spliced in HCMV	1
U85	U85	U85		Probable membrane glycoprotein, related to OX-2	1,2
U86	U86	U86	UL122	IE-A protein, HCMV IE2 homology	1,2 1
U89	U90 ex3	U89		IE-A transactivator	1,2
U90	U90 ex2	U90		IE-A transactivator	1,2
U91	U91	U91		Probable membrane glycoprotein	2
H7				Unlikely to encode protein	2
U95	U95	U95		US22 gene family, MCMV IE2 homolog	1,2 1
H8	U100 ex10			Envelope glycoprotein gp105	2
U98	U100 ex3	U98		Envelope glycoprotein gp105	2
U99	U100 ex2	U99		Homology to HHV-6 gp82/105	1,2
U100	U100 ex1	U100		Homology to HHV-6 gp82/105	1,2
H1'				Unlikely to encode protein	2
DR1'	DR1' ex1	DR1'		US22 gene family DR1/DR6 homology	1,2 1
DR2'	DR1' ex2	DR2'		US22 gene family	1,2
H2'				Unlikely to encode protein	2
DR6'	DR6' ex1	DR6'		US22 gene family DR1/DR6 homology	1,2 1

HHV-7 ORF (strain [a])		Homologues [b]		Predicted function	Ref [c]
(JI)	(RK)	HHV-6	HCMV		
DR7'	DR6' ex2	DR7'		US22 gene family Transactivator	1,2 1
H3'				Unlikely to encode protein	2
H4'				Unlikely to encode protein	2

[a] HHV-7 strain JI, Nicholas *et al* 1996; RK, Megaw *et al* 1998

[b] HHV-6, Gompels *et al* 1995; HCMV, Chee *et al* 1990

[c] Reference 1, Nicholas *et al* 1996; 2, Megaw *et al* 1998; 3, Menegazzi *et al* 1999; 4, Pellet and Dominguez 2001

1.1.5 Gene expression

Lytic expression of HHV-7 genes appears to follow the temporal regulated pattern (Menegazzi et al., 1999), common to the herpesviruses (Figure 1.3). Herpesvirus genes can be classified into immediate early (IE or α), early (E or β) or late (L or γ) genes, based on their patterns of regulated expression (Chambers et al., 1999; Honess and Roizman, 1974; Mirandola et al., 1998; Wagner et al., 1972). IE genes are expressed immediately-early during infection independently of *de novo* protein synthesis. They provide transcription regulatory functions for the E genes. Transcription of the E genes rely on the IE genes but not viral DNA replication, and their products are often associated with viral DNA replication functions. L genes often encode structural proteins, and their expression is almost completely dependent on viral DNA replication. A few of the HHV-7 genes have been classified into their kinetic classes using inhibitory drugs that block protein synthesis (cycloheximide and emetine) and DNA replication (phosphonoacetic acid) (Menegazzi et al., 1999). More details on HHV-7 and herpesvirus gene expression are given in Chapter 5.

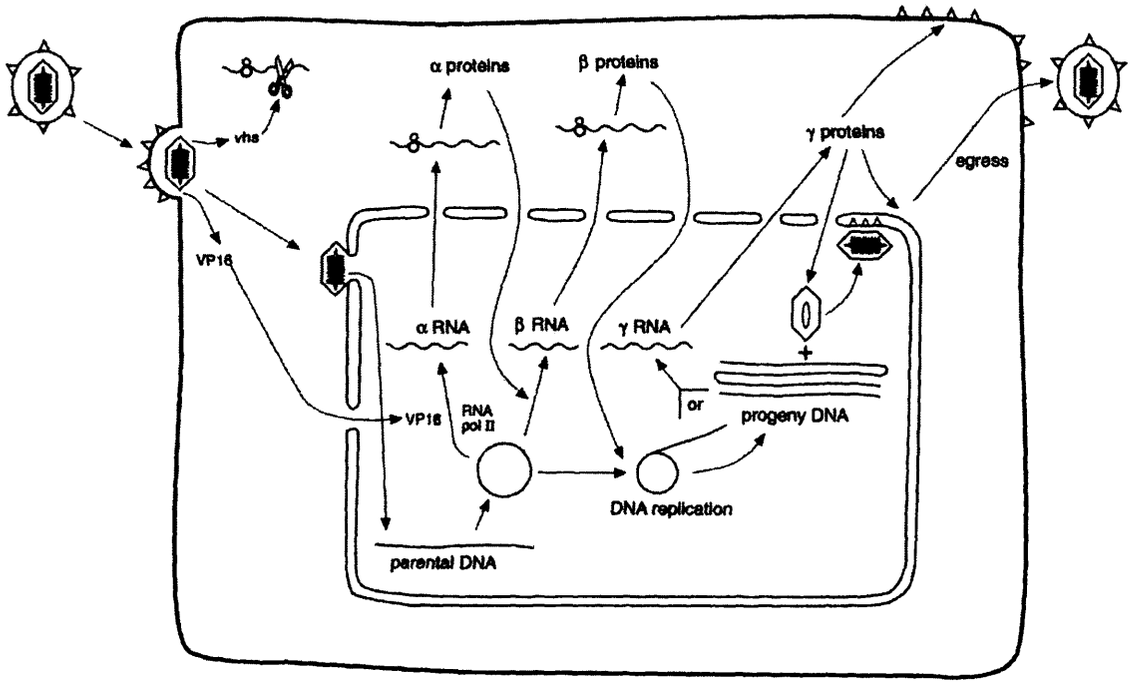


Figure 1.3 Temporal regulation of herpes simplex virus (HSV) gene expression. After virion binding, internalisation and nuclear entry, the viral DNA circularizes and is transcribed by host RNA polymerase II to give first the IE or α mRNAs. Five of the six HSV IE proteins act to regulate viral gene expression in the nucleus. They transactivate E or β gene transcription. The E proteins are involved in replicating the viral DNA molecule. Viral DNA synthesis stimulates L or γ gene expression. The L proteins are involved in assembling the capsid in the nucleus and modifying the membranes for virion formation. The mature enveloped virion exits the cell. Tegument proteins vhs (virion host shutoff) and VP16 are involved in degrading mRNA and activating IE gene expression respectively. (Reproduced from Roizman and Knipe, 2001)

1.1.6 Cell entry and Receptor

The cell surface CD4 molecule is believed to be the receptor for HHV-7 infection, supported by various experimental findings. It has been demonstrated that, when several non-susceptible cell lines, Jurkat (T-cell line), U937 (monocyte), SAS413 and K562 (myelocytes) were transfected with CD4, they became susceptible to HHV-7 infection, as measured by cell morphology and indirect immunofluorescence (Yasukawa et al., 1997; Zhang et al., 2000). Radiolabelled HHV-7 virus binds specifically to CD4-transfected HeLa cells but not to unmodified HeLa cells (Lusso et al., 1994). HHV-7 infection was inhibited in a dose-dependent manner by preincubation with soluble CD4, anti-CD4 monoclonal antibodies, or gp120, the viral ligand for CD4 in human immunodeficiency virus (HIV) infection of $CD4^+$ T cells (Lusso et al., 1994; Yasukawa et al., 1997). Flow cytometry studies showed that surface CD4 is reduced in HHV-7-infected PBMCs and SupT1 cells, while other T-cell markers CD2, CD3 and CD44 remained unchanged (Lusso et al., 1994; Zhang et al., 2000). Downregulation of receptor molecules (CD4 and CCR5) during infection has also been observed in HIV (Piguet et al., 1999), possibly as a viral strategy to prevent superinfection (Little et al., 1994; Mitchell et al., 2003), and is believed to play an important role in HIV replication (Levesque et al., 2004; Lundquist et al., 2002). The above experimental results therefore conclude that CD4 is essential for HHV-7 infection of $CD4^+$ T cells.

To date, the viral ligand for CD4 has not been identified, and the binding site on CD4 has not been mapped. However, clues can be obtained by comparing HHV-7 and HIV-1 interaction with CD4. HHV-7 infection is blocked by a monoclonal antibody which binds to the V3/V4 domains of CD4, as well as by those that bind to the V1 region (Lusso et al., 1994); HIV-gp120 binding to CD4 is via the V1 domain, unaffected by anti-V3/V4 antibodies. Moreover, pre-exposure of live or UV-inactivated HHV-7 inhibits HIV-1 infection in macrophages (Crowley et al., 1996). Therefore the HHV-7 binding site on CD4 is probably different to that of HIV-1, but may span multiple domains so as to block HIV-1 binding. Understanding HHV-7/CD4 interactions may provide clues for further antiviral strategies against HIV-1.

Although CD4 is essential for HHV-7 infection, many $CD4^+$ T-cell lines are either only weakly or not at all susceptible to the virus (Yasukawa et al., 1997), suggesting that CD4 alone is not sufficient for HHV-7 infection. The search for an alternative or co-receptor led to the identification of cell surface proteoglycans as potential candidates. Heparin, the natural ligand of heparan sulphate proteoglycans, inhibits HHV-7 binding to SupT1, formation of syncytia and viral replication (Secchiero et al., 1997c), suggesting heparan sulphate may play a role in HHV-7 infection. Several HHV-7 proteins have been proposed as viral ligands for attachment to cell surface proteoglycans, including glycoprotein B encoded by U39 (Secchiero et al., 1997c) and gp65 encoded by U98/100 (Skrincosky et al., 2000). Heparan sulphate proteoglycans have been reported to be the cellular receptors for alphaherpesviruses (Spear and Longnecker, 2003), herpes simplex virus (Campadelli-Fiume et al., 2000), CMV (Compton et al., 1993), VZV (Zhu et al., 1995), and Papillomavirus (Shafti-Keramat et al., 2003).

The role of CXCR4 as a co-receptor has also been studied with inconclusive results. Two studies reported downregulation of surface CXCR4 during HHV-7 infection (Secchiero et al., 1998c; Yasukawa et al., 1999), whereas another study showed that CXCR4 transcription is not modified by HHV-7 and does not act as a co-receptor (Zhang et al., 2000). More research is needed to strengthen our understanding of the molecular interactions during HHV-7 entry of $CD4^+$ T cells.

1.1.7 Cell tropism & Persistence

HHV-7 has a restricted cell tropism of T-lymphocytes *in vitro*, demonstrated by double immunofluorescence of infected cord blood mononuclear cells (CBMCs) (Berneman et al., 1992a). SupT1, a $CD4^+$ T-lymphoblast cell line derived from non-Hodgkin's T cell lymphoma, is the only cell line shown to support sustained, productive HHV-7 infection *in vitro* (Ablashi et al., 1998; Berneman et al., 1992a), shown by CPE observation, IFA for viral antigen, and electron microscopy. On the other hand, a number of cell lines are shown to be non-susceptible to HHV-7 infection (no viral antigen by IFA and no viral DNA by PCR), including several $CD4^+$ and $CD4^-$ T-cell lines HSB2, MoLT-3, J-Jhan, MoLT-4, Jurkat (Ablashi et al., 1998); HUT-78, MT-4, (Zhang et al., 2000); EU-T-01, Molt-15, HO8-HT (Yasukawa et al., 1997); B-cell lines Daudi, Raji, MY-LCL, MJ-LCL (Yasukawa et al., 1997); and macrophages (Crowley et al., 1996). In contrast to Crowley et al, it has also been reported that HHV-7 is able to establish productive infection in macrophages up to 2 weeks post inoculation, while viral DNA could be detected for up to a month but not viral mRNA, therefore HHV-7 may establish latent infection in macrophages (Zhang et al., 2001a).

The *in vivo* reservoirs of HHV-7 are not known, but the virus is frequently isolated from peripheral blood (Fernandez et al., 2002) and saliva (Wyatt and Frenkel, 1992). Viral DNA is present in 55 to 100% of salivary gland specimens (Sada et al., 1996; Yadav et al., 1997). HHV-7 has also been detected by immunohistochemistry in other tissues including lung (7/8), skin (7/11), mammary glands (4/7), liver (5/8), kidney (8/10) and tonsils (2/6) (Fujisaki et al., 1998; Kempf et al., 1998). However, only a small proportion of the these samples (lung, 1/8; skin, 2/11; and mammary glands, 2/7) also scored positive for viral DNA in PCR, compared to salivary glands (6/10). Tissue samples of brain, large intestine and spleen were HHV-7-negative.

1.1.8 Transmission & Prevalence

HHV-7 is a highly prevalent virus, with over 90% adults being seropositive to the virus (Clark et al., 1993; Wyatt et al., 1991). Primary infection is thought to occur in early

childhood. Some data showed that, like HHV-6, antibodies to HHV-7 were detected in infants under 2 years old (Clark et al., 1993), whereas others suggest a later seroconversion in 3 to 4 year olds (Wyatt et al., 1991). Another study reported HHV-7 prevalence peaks at 60 % in 11-13 year olds (Yoshikawa et al., 1993). Infectious HHV-7 virus has been readily isolated from saliva of healthy adults (Wyatt and Frenkel, 1992; Yoshikawa et al., 1993), suggesting the saliva plays a role in spreading the virus. It has been shown that viral DNA isolated from the saliva of children shared similar restriction patterns to their parents, providing further proof for the horizontal transmission of HHV-7 via close contact between parent and child (Takahashi 1997). In addition, breast feeding may also be a possible transmission route, as viral DNA has been detected in breast milk, and the seropositivity rate is slightly higher in children who are breast-fed than bottle-fed (Fujisaki et al., 1998).

1.1.9 Disease association & Immune response

The ubiquitous nature of HHV-7 makes it difficult to identify disease associations. Studies have linked the virus to exanthem subitum (ES) (Tanaka-Taya et al., 2000; Torigoe et al., 1995; Ueda et al., 1994), a childhood illness characterised by fever and skin rash, possibly because HHV-7 reactivates HHV-6 (Tanaka-Taya et al., 2000), the primary causal agent of ES (Yamanishi et al., 1988). There are controversial views on the role of HHV-7 in pityriasis rosea (PR, a relapsing rash). One study detected HHV-7 DNA in plasma, PBMC and skin samples from patients with PR, but not in controls (Drago et al., 1997), whereas another study concluded there was no significant difference between the prevalence of HHV-7 in PR skin lesion samples and control samples, either by nested PCR for HHV-7 DNA, or by IFA for antigen (Kempf et al., 1999). HHV-7 may be an opportunistic pathogen responsible for transplant complications, for example delayed neutrophil engraftment in bone marrow recipients (Chan et al., 1997). It may also have associations with CMV disease in liver (Mendez et al., 2001), kidney (Osman et al., 1996) and bone marrow (Chan et al., 1997) transplant patients. At least some of these correlated with HHV-7 viraemia and/or DNA load (Brennan et al., 2000; Chan et al., 1997; Mendez et al., 2001). There are no significant links between HHV-7 and chronic fatigue syndrome (CFS) (Ablashi et al., 1998), and the isolation of HHV-7 from CFS patients may be the consequence of CFS

rather than cause (Secchiero et al., 1994). In summary, to date no disease has been confirmed to be caused by HHV-7, with the virus apparently very well adapted to persistence and transmission in the human host.

There are very few published studies on the host immune response to HHV-7, possibly due to the virus's lack of pathogenicity. However, information can be gathered from the data of serological prevelance studies. Primary infection by HHV-7 leads to the production of HHV-7–specific antibodies in 40% of healthy children by 44 months (Wyatt et al., 1991). Antibody titres to HHV-7 in children are high in the first 2 months (94.4%), possibly due to presence of maternal antibodies, followed by a decline between 6-7 months (30%). A rise in HHV-7 antibodies after 8 months (Clark et al., 1993) was possibly due to the infected infants producing their own antibodies to the virus after the reduction in maternal antibodies. These antibodies are apparently protective against the virus, as sera from healthy individuals and CFS patients who are HHV-7–positive inhibit HHV-7–induced CPE *in vitro* (Secchiero et al., 1994). HHV-7 antibodies are sustained into adulthood, as sera from 24 out of 26 (92%) healthy adults are seropositive for HHV-7 (Wyatt et al., 1991).

In addition to a humoral response, it has also been observed that $CD4^+$ T-cells produce gamma interferon (IFNγ) and show a proliferative response to HHV-7 antigen (Yasukawa et al., 1993). Specificity tests showed that 28% of the T-cell clones reacted against both HHV-6 and HHV-7. Since primary infection by HHV-7 seem to occur after HHV-6 (Wyatt et al., 1991), perhaps the immune system is already equipped for protection against HHV-7 upon primary infection. In addition, during latency, HHV-7 probably evades the immune system by minimising viral gene expression (Menegazzi et al., 1999).

1.1.10 Cellular interaction

It would not be surprising that some of the HHV-7 gene products interact with the host. HHV-7 induces a range of effects in infected cells. Examples include the downregulation of the cellular receptor CD4 (Furukawa et al., 1994; Lusso et al., 1994; Secchiero et al., 1997b) and downregulation of CXCR4 (Secchiero et al., 1998c).

HHV-7 infection of $CD4^+$ T cells also dysregulates expression and activity of cyclin B and the cyclin-dependent kinase cdc2, leading to cell cycle arrest and polyploidisation of infected cells (Secchiero et al., 1998a). These interactions may be involved in apoptosis regulation and virion production (Secchiero et al., 1998b). A possible immune evasion mechanism has also been observed; HHV-7 infected cells had increased levels of Tumour necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) but decreased surface TRAIL receptor TRAIL-R1. As a result while bystander cells were killed by TRAIL-mediated cytotoxicity, HHV-7 infected cells remained resistant (Secchiero et al., 2001). The underlying molecular mechanisms of these effects are not characterised, and the participation of any HHV-7 genes or proteins remains unknown. Although the HHV-7 genome has not been extensively studied, information can be deduced from their homologues in other herpesviruses, especially HHV-6 and HCMV. Genes in the HHV-6 IE-A and IE-B loci, DR6, U3 and U27 are reported to have transcription regulatory functions. It has been shown that viral promoters such as HIV-1 LTR (Geng et al., 1992) and EBV EBNA-2 (Cuomo et al., 1998), in addition to the cellular promoter for CXCR4 (Yasukawa et al., 1999) are affected by HHV-6 regulators. HHV-7 U39 encodes the glycoprotein gB, which binds to cell surface heparin sulphate proteoglycans (Secchiero et al., 1997c), and may play a role in attachment and penetration, as in HHV-6 (Takeda et al., 1996). Glycoproteins gH and gL, encoded by U48 and U82 respectively, may be involved in membrane fusion, at least in HHV-6 (Qian et al., 1993). HHV-6 has also been reported to simultaneously stimulate host protein synthesis and shut down DNA synthesis (Pellett, 2001), thus it is possible that HHV-7 may induce changes of similar scale in the host during infection.

1.1.11 Life cycle

The life cycle of HHV-7 has not been extensively studied, but can be deduced from other properties of the virus, as mentioned in the previous sections. It is widely accepted that HHV-7, like all herpesviruses, persists in the host by a combination of lytic and latent infection, possibly in various sites or tissue types (Black and Pellett, 1999).

1.1.11.1 Latency/Persistence

After acute infection, the virus probably establishes latency in T-cells, as no expression of viral lytic genes were detected by RT-PCR in PBMCs isolated from patients despite having a high viral load (Menegazzi et al., 1999). Moreover, virus can be isolated from activated but not from resting PBLs, suggesting the virus may be reactivated from latency under T-cell activating conditions (Frenkel et al., 1990). Perhaps HHV-7 from activated T-cells is disseminated in the host via blood (Fernandez et al., 2002), and establishes a persistent infection in various other tissues, where viral DNA has been detected (section 1.1.7). Finally, HHV-7 persistence may be maintained by lytic production in salivary glands, as the virus is readily shed in saliva (Kempf et al., 1998; Wyatt and Frenkel, 1992; Yoshikawa et al., 1993).

1.1.11.2 Lytic infection

Lytic infection begins with reactivation or with viral attachment to specific surface receptors and penetration of host cells (section 1.1.6). HHV-7 encodes the necessary genes for replication and virion assembly (section 1.1.4), and the regulated expression of these genes (section 1.1.5) drives the production cycle of the virus. Electron microscopy (Black et al., 1997) detected DNA-containing capsids in the nucleus 3 days after infection. Capsids appear to acquire tegument in the perinuclear space, seen as a pool of electron-dense material surrounding the capsid. Tegumented capsids can be observed in the cytoplasm, and probably acquire the envelope by budding upon the golgi membrane. Matured virions are released into extracellular space after 5 days, causing cell death by apoptosis and necrosis *in vitro* (section 1.1.3).

In conclusion, HHV-7 is an efficient virus, establishing persistent infection in most of its human host population, whilst causing minimal disease, avoiding host immune response, and maintaining the ability to replicate. Studying HHV-7 may therefore reveal more detail about how the human host can tolerate the presence of commensal viruses.

1.2 DNA Microarrays

1.2.1 The rise of the genome, transcriptome, and proteome

In recent years, the advantages in studying gene expression using microarray technology have been widely acknowledged (Harrington et al., 2000; Lockhart and Winzeler, 2000). This technology excelled due to its high throughput nature, reproducibility and relative ease of use. The expansion of microarray technology can also be attributed to its parallel development with three technical advancements in biomedical science in the recent years. These are the rapid generation of genomic data from various genome projects, development in automated precision robotics and the improvement in computational data management.

To date, the complete genome sequences of many organisms have been elucidated, including the human genome (McPherson et al., 2001). Genomes of many pathogens have been sequenced, including 217 bacterial genomes (http://www.tigr.org), 799 virus genomes (http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/10239.html), as well as many protozoan, fungus and helminth species (http://www.sanger.ac.uk). A large number of genome projects are currently ongoing. Many of the completed genome sequences have been annotated for potential sites of expressed open reading frames (ORFs), transcription initiation site, promoter sequences and splicing information. However, genome sequence data alone do not provide information about what the genes do (prediction by sequence homology is limited to the number of genes of which the functions we already know), how their expression is regulated, or what roles they play in the biological processes in coordination with other genes. The conventional methods for gene analysis do not have the capacity to match the vast quantity of data generated, hence there is a need for large scale analysis of genome data. Microarray technology provides a bridge for this gap. This high throughput technology allows many genes to be analysed simultaneously and rapidly.

Early arrays contained only several hundred elements printed on nylon membranes. The production of high density solid substrate microarrays relied upon automated precision robotics technology to deposit tens of thousands of elements onto the substrate

surface (DeRisi et al., 1997; Schena et al., 1995). Therefore the advancement in precise and high density robotic deposition facilitated the development of microarray technology.

Good data management systems are required for many stages in microarray experiments (Chicurel, 2002). First of all, accurate genomic data available in genome databases have facilitated the production of array elements of specific sequences (Kellam, 2001). Annotated genomic databases have also been developed, for example, VIDA is a specialist database for viruses (Alba et al., 2001) that organise ORFs from many complete viral genome sequences into homologous protein families. A set of standards - Minimal Information About a Microarray Experiment (MIAME) (Brazma et al., 2001) has been created for the submission of microarray results to the public repository (http://www.ebi.ac.uk/arrayexpress/), providing a interface for scientists to access and compare microarray data. These computational advances therefore played important roles in the development of high-density microarrays.

The high throughput nature of microarrays made them ideal for genome-wide analysis applications. Comparative genome hybridisation (CGH) is a term given for using microarrays in genotyping (Hacia, 1999; Stoughton, 2005) and single nucleotide polymorphism (SNP) studies (Shen et al., 2005). CGH arrays have been developed and extensively applied to investigate genetic diversity in a number of bacterial systems including *Helicobacter pylori*, *Campylobacter jejuni*, *Yersinia pestis*, *Mycobacterium tuberculosis*, and *Streptococcus pneumoniae* (Hakenbeck et al., 2001; Hinchliffe et al., 2003; Kato-Maeda et al., 2001; Salama et al., 2000). In a study of 15 virulent *Helicobacter pylori* strains, 362 out of 1643 genes analyzed were identified as strain-specific (Hakenbeck et al., 2001; Hinchliffe et al., 2003; Kato-Maeda et al., 2001; Salama et al., 2000). A similar genotyping analysis comparing 22 strains of *Yersinia pestis* and 10 strains of *Yersinia pseudotuberculosis* identified a number of differences that may relate to the evolutionary pressures on these species in their respective niches (Hinchliffe et al., 2003). An oligonucleotide array containing 82 polymorphic oligonucleotides of the 16S ribosomal RNA gene has been used for genotyping of polymorphisms and identification of mycobacterium species (Troesch et al., 1999). CGH arrays are not limited to bacterial studies. DNA microarrays have been made for genotyping of the amoeba parasite *Entamoeba histolytica* (MacFarlane et al., 2005a),

and to identify deletions in BCG vaccine strains (Behr et al., 1999). Wang *et al* developed an oligonucleotide (70-mer) microarray to detect 140 different viruses using known sequence data (Wang et al., 2002a).

Perhaps one of the most widely used applications of microarrays is the monitoring of gene expression (section 1.2.4). By means of measuring abundance of all mRNA transcripts in a sample simultaneously, microarrays provide a global picture of gene expression pattern (the transcriptome), which shows not only the expression of many genes individually, but also the coordinated expression pattern of the different genes at a given condition. This often reveals the physiological state of the system, such as stages of cell or tissue development (Pang et al., 2005), or responses to external stimuli such as drug treatments (Gerhold et al., 2001; Gray et al., 1998; Scherf et al., 2000). Transcription studies have helped towards understanding the mechanisms of many diseases, including many cancers such as Hodgkin's lymphoma (Cossman, 2001), hepatocellular carcinomas (Graveel et al., 2001; Okabe et al., 2001), breast cancer (Murphy et al., 2005), colorectal carcinogenesis (Kitahara et al., 2001). However, the accessibility of tumour tissues presents a challenge to such microarray studies, as cancerous cells need to be separated from normal cells in order to give differentiated expression patterns. For example, Kitahara *et al* used microarrays to compare the transcription profiles of colorectal cancer cells to noncancerous colonic epithelia, separating them using laser-capture microdissection, and amplifying the RNA signal. The differential expression pattern observed indicated that certain processes were accelerated in the cancer cells (metabolism, production of reactive oxygen species), whereas regulation of cell cycle, transcription and mitosis were impared, compared to normal tissues. Microarrays have also offered insights into microbial infections (section 1.2.4) and facilitated identification of disease subphenotypes and diagnosis (Helmberg, 2001; Petrik, 2001). The diversity of microarray applications in gene expression profiling has been extensively reviewed (Heller, 2002; Stoughton, 2005).

Based on the observation that genes with similar function are often co-expressed (Bassett et al., 1999), gene expression profiling can be applied to predict functions of unknown genes that are expressed with characterised ones, facilitating the annotation of microbial genome, for example, in *Saccaromyces cerevisiae* (Eisen et al., 1998). More

details of microarray-based gene expression analysis of microorganisms will be given in section 1.2.4.

Gene expression experiments provide valuable information about the global pattern of transcription in a biological system, but they do not address post-translational modifications of the protein product. Various methods are used in proteome studies, such as mass spectrometry, 2-dimentional (2d-) gel electrophoresis and protein chips (Pandey and Mann, 2000). Proteomic mass spectrometry analysis of *Plasmodium falciparum*, the protozoan responsible for malaria, has been used to characterise 2400 proteins at four stages of the protozoan's life cycle (Florens et al., 2002). Protein chips containing 35 – 149 protein probes have been developed for several bacteria, such as *Salmonella enterica* (Cai et al., 2005), *Helicobacter pylori* (Das et al., 2005), and *Neisseria meningitides* (Steller et al., 2005). A protein chip has also been produced for yeast, which contains purified proteins over-expressed from 5800 open reading frames of the yeast genome (Zhu et al., 2001). Miller et al. used protein arrays with cellular genes to identify potential prostate cancer biomarkers in prostate cancer serum (Miller et al., 2003). However, although proteomics targets the 'final product' of gene expression, they are more difficult to conduct, and possibly less robust than DNA microarrays (Lockhart and Winzeler, 2000). The 2d-gel approach is limited by dynamic range hence only efficient for the analysis of a subset of proteins (Pandey and Mann, 2000). Production of protein chips is more complex and costly than DNA microarrays, because proteins have a high variety of binding chemistries, lower stability, and often require multimerization and post-translational modifications. A lack of simple amplification processes like PCR also makes it more difficult to generate large quantities of protein (LaBaer and Ramachandran, 2005). Therefore currently protein arrays are of smaller scale and diversity than the DNA microarrays.

1.2.2 Principles of DNA microarray technology

Microarray technology is a large-scale competitive hybridisation technique that measures the relative abundance of many DNA or RNA species in parallel. A microarray essentially contains a large set of nucleic acid probes of known sequences immobilised onto a solid support, known as the substrate. The sample in question (a mixed population of DNA or RNA) is labelled, usually with fluorescence or radioactivity, and applied to the microarray. The different species of DNA or RNA in the sample hybridise to their corresponding array probes of complementary sequences. The detected fluorescence or radioactivity of each hybridised probe is a measure of the relative abundance of that particular gene in the sample. There are many types of microarrays and experimental design varies, but all share the common principle of detecting abundance of many nucleic acids simultaneously using a large number of probes.

1.2.2.1 Microarray platforms

There are three main types (or platforms) of microarrays, based on substrate material and the nature of probes. They differ in a range of features but for simplicity we will call them membrane arrays, spotted arrays and oligonucleotide chips in this thesis. Membrane arrays are the earliest form of microarray, using nylon membrane as the substrate on which only a few hundred DNA probes (PCR product or clones of a expression library) are deposited (Ahn et al., 2002; Jenner et al., 2001). Membrane arrays are usually used with radioactive-labelled samples. On the other hand, both spotted arrays and oligonucleotide chips use glass as substrate, which allows for more densely-arranged probes thus have a higher capacity (10,000s of probes per glass slide) than membrane arrays. Spotted arrays and oligonucleotide chips are both used to detect fluorescent-labelled samples but they differ in the nature of probes and method of probe generation. Spotted arrays are made in a similar manner to membrane arrays. DNA probes are pre-made (by cloning or by PCR) and then deposited or 'spotted' onto the substrate. The probes are generally longer (200-2000 bases), although oligonucleotide probes as small as 50-mer has been described (Kane et al., 2000). In contrast,

oligonucleotide chips are made by *in situ* synthesis of oligonucleotide probes 25-mer to 60-mer in length, and usually more than one probe per gene are present on the array (Harrington et al., 2000).

The presence of multiple probes per gene gives oligonucleotide chips the ability to detect splice variants and to cover the entire genome, rather than just the predicted open reading frames (Lockhart et al., 1996; Pease et al., 1994). However, generating oligonucleotide probes *in situ* is costly, time consuming, and can result in products of variable yield. On the other hand, deposition of probes onto spotted arrays is faster, cheaper and can be easily implemented in non-commercial research laboratories, as in the case of many viral arrays (Chambers et al., 1999; Paulose-Murphy et al., 2001; Stingley et al., 2000). It has been shown that the sensitivity of a 50-mer oligonucleotide chip was comparable to that of spotted arrays with longer probes, but some of these 50-mer probes cross-hybridised to sequences that share a 75-80% similarity with the target sequence (Kane et al., 2000), suggestive that the shorter probes may have lower specificity. To improve specificity, short oligonucleotide chips normally contain mismatched sequences as controls (Lockhart et al., 1996). Nonetheless, some laboratories have created spotted arrays with shorter oligonucleotide probes (Chambers et al., 1999; Kane et al., 2000; Okamoto et al., 2000; Stingley et al., 2000), demonstrating the convenience of spotted arrays and the value of shorter probes, especially for applications that require multiple probes per gene, such as single nucleotide polymorphism studies (Okamoto et al., 2000). Many studies have shown that specificity and sensitivity increase with probe size and the optimal probe length is 60 – 70-mer, with only modest increases above this size.

A number of studies have compared microarray data generated using different array platforms with contradictory conclusions. Some publications showed that data obtained with different array platforms were agreeable (Kane et al., 2000; Yuen et al., 2002); others reported little correlation between platforms (Kothapalli et al., 2002). However, recent studies argued that when care was taken to compare relative expression rather than absolute values, the patterns of expression were similar on both platforms (oligonucleotide chips and spotted arrays) even though the relative amplitude of change was greater in one platform than the other (Irizarry et al., 2005; Larkin et al., 2005). It was also reported that different array platforms actually had *less* effect on the observed

expression variations than both biological differences of the samples (Larkin et al., 2005) and the laboratories in which the experiments were conducted (Irizarry et al., 2005). When discrepancies were observed between oligonucleotide chips and spotted microarrays, quantitative RT-PCR results supported neither platform but gave a third pattern (Larkin et al., 2005), suggesting neither platform was more accurate that the other, and the discrepancies might have been due to splice variants. These findings concluded that the choice of array platform will not have significant effect on the accuracy of data obtained, providing careful data analysis was performed. For our study we have chosen to deposit PCR-amplified DNA probes (200 – 2000 bases) onto spotted glass microarrays for their higher density than membrane arrays, also for the relatively low-cost production that is easy to set up in-house.

1.2.2.2 Microarray and gene expression analysis

The use of microarrays in the monitoring of gene expression is based on measuring transcript (mRNA) abundance. This requires extraction of total or mRNA from the sample of interest, which are then reverse transcribed into cDNA. The cDNA is labelled with fluorophors either directly during reverse transcription or by indirect linkage afterwards. For spotted arrays it is most common to use the two-colour labelling method, in which two cDNA samples are labelled as a pair, sample 1 with the red fluorescent Cy5 dye and sample 2 with the green fluorescent Cy3 dye. The pair are then mixed and co-hybridised onto one array. The array is then scanned and the fluorescence intensities measured at around 635 nm (Cy5) and 530 nm (Cy3) of each hybridised probe represent the relative abundance of the particular mRNA sequence in sample 1 and 2, respectively (Figure 1.4). This two-colour system is a very important aspect of the spotted microarray. Microarray experiments are often designed to investigate changes in gene expression over a time series (Ahn et al., 2002; Jenner et al., 2001), under different conditions, e.g. drug treatment vs untreated (Wilson et al., 1999), or to compare expression between two types of cells, e.g. normal versus tumour (Okabe et al., 2001). All of these require the use of more than one microarray and subsequently, comparing data from different arrays. Such comparison is subject to errors as each array hybridisation is unique and variations occur in many processes, such as the amount of DNA probe per spot, labelling efficiency and hybridisation efficiency

(Quackenbush, 2002). Such errors can be reduced using the two-colour system. Because each array probe is hybridised to two samples with different coloured labels, any array-based variations should have the same effect on both Cy5 and Cy3 samples. Therefore comparing the Cy5/Cy3 ratios of each probe across different arrays allows elimination of these errors. This can be achieved by labelling the common reference RNA with one colour (usually the Cy3) and the RNA of interest in the other (Cy5). The same Cy3-labelled reference is used to hybridise all arrays, hence acting as a control so that the Cy5/Cy3 ratios of one gene on different arrays can then be compared.

There are several options of binding chemistry for probe attachment on glass slides, including aminosilane-coating, epis-coating, conjugated coating. Hybridisation depends on the type of array and the binding chemistry (Zammatteo et al., 2002). Spotted glass arrays generally require denaturation and blocking prior to hybridisation to prevent non-specific binding. During the actual hybridisation, optimal temperature, time, and hybridisation buffer containing the correct ingredients are essential to optimise specific binding. The conditions for post-hybridisation washing are also important in order to maximise signal-to-background ratios. Hybridised arrays are scanned under lasers operating at around 635 and 530 nm to excite the fluorophors to produce fluorescent emissions, which are amplified by photomultiplier and recorded as images. These images are then analysed by imaging softwares to produce numerical data (Hegde et al., 2000) for data processing and analysis, discussed in the next section.

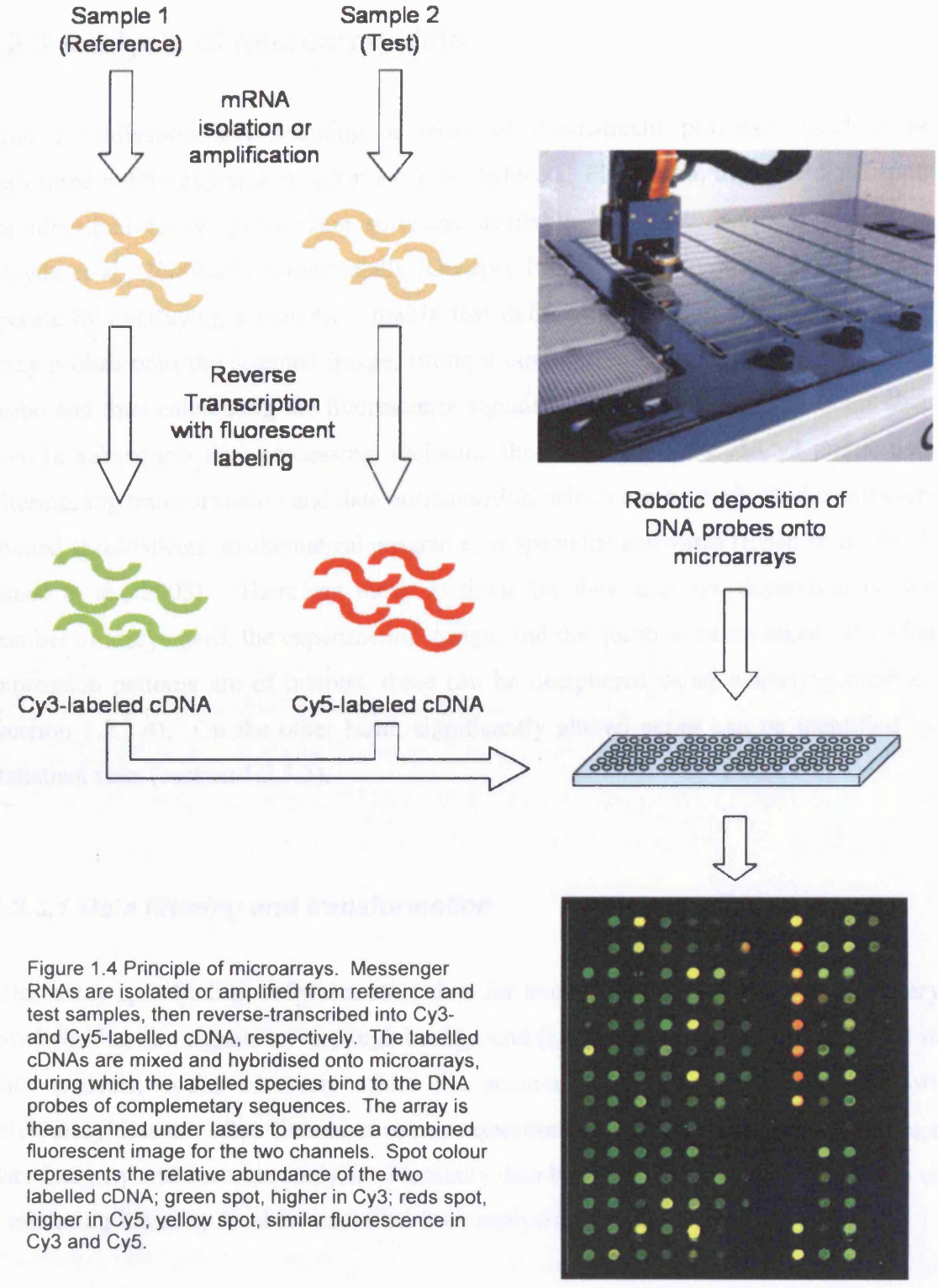


Figure 1.4 Principle of microarrays. Messenger RNAs are isolated or amplified from reference and test samples, then reverse-transcribed into Cy3- and Cy5-labelled cDNA, respectively. The labelled cDNAs are mixed and hybridised onto microarrays, during which the labelled species bind to the DNA probes of complemetary sequences. The array is then scanned under lasers to produce a combined fluorescent image for the two channels. Spot colour represents the relative abundance of the two labelled cDNA; green spot, higher in Cy3; reds spot, higher in Cy5, yellow spot, similar fluorescence in Cy3 and Cy5.

1.2.3 Analysis of microarray data

After hybridisation and scanning, a series of downstream processes need to be performed before expression patterns can be deduced. First of all, the hybridised spots are identified using spot-finding software, available both open-source reviewed in (Hegde et al., 2000) and commercially (Genepix Pro, Axon Instruments, USA). They operate by overlaying a numerical matrix that defines the position and identity of the array probes onto the scanned image, fitting a circlular 'spot' around each hybridised probe and then calculating the fluorescence signals of every spot. These data are then used in subsequent data processing, including the calculation of Cy5/Cy3 ratios, data filtering, log transformation and data normalisation, which can be performed using user-created spreadsheets, mathematical programs, or specialist softwares (Eisen et al., 1998; Saeed et al., 2003). There are many methods for data analysis, depending on the number of arrays used, the experimental design, and the question being asked. If global expression patterns are of interest, these can be deciphered using clustering methods (section 1.2.3.4). On the other hand, significantly altered genes can be identified by statistical tests (section 1.2.3.3).

1.2.3.1 Data filtering and transformation

Microarray spot-finding softwares store data for every spot, including those with very low hybridisation signals or very high background (i.e. noisy data), which if included in data analysis, could adversely affect the accuracy of interpretations. Therefore microarray data are often filtered to remove questionable or low-quality data, to ensure data integrity and accurate analysis. Normally data below a signal-to-noise threshold or a minimum intensity level are excluded from analysis.

The Cy5/Cy3 ratio is a useful way of representing data but it gives a prejudiced view of up-regulated and down-regulated genes. Genes with double the signal in Cy5 compared to Cy3 have a ratio of 2, whereas genes with twice as much fluorescent in Cy3 than Cy5 have a ratio of 0.5, despite the fold change is the same. When $\log_2$ transformed these ratios appear as +1 and -1, respectively, hence presenting up- and down-regulated genes

by the same magnitude fold-change. Log_2 transformation also converts the spread of data to normal distribution, thus allowing the use of parametric statistical methods on the data. It is a common pre-analysis adjustment of microarray data (Quackenbush, 2002).

1.2.3.2 Data normalisation

Microarray data need to be normalised before interpretation. The purpose for normalisation is to adjust the data so that the effects of experimental procedures on the data are removed as much as possible. For example, the efficiency of labelling and hybridisation of the two channels can never be exactly identical, resulting in systematic errors in the Cy5/Cy3 ratios. There are several widely used methods to reduce this dye-bias effect, based on total intensity, regression, and ratio-statistics. All three approaches assume that the average Cy5/Cy3 expression ratio of all (or a subset of control) genes on the array should be equal to one. Total intensity normalisation assumes the total intensity of all (or a subset of) spots in one channel is equal to that of the other channel. A correcting factor is applied to each spot, so that the total $\log_2 Cy5_{total}/Cy3_{total}$ ratio is equal to zero (Quackenbush, 2002). Regression-based normalisation adjusts the intensities so that the best-fit slope of the Cy5/Cy3 scatter plot is equal to one (Hedenfalk et al., 2001). Ratio statistics-based normalisation is an iterative method that adjusts Cy5/Cy3 ratios to normalise the mean Cy5/Cy3 ratio to one (Chen et al., 2002).

Another reason for data normalisation is the phenomenon of 'intensity-dependent local variation'. It has been observed that the $\log_2$ ratios of low-intensity data generally deviate further from zero than the high-intensity spots. As a result these genes may be falsely classified as up- or down-regulated genes. To understand this phenomenon consider a high-intensity gene spot H and a low-intensity gene spot L. If gene H gave intensity signals of 2200 and 2000 in Cy5 and Cy3 channels respectively, the $\log_2$ Cy5/Cy3 ratio is +0.13 and the expression of gene H would be regarded as relatively unchanged. On the other hand, gene L with signals of 400 and 200 will have a $\log_2$ Cy5/Cy3 ratio of +1, and appear as an up-regulated gene, even though both genes have a small difference of 200 intensity units between the two channels. This intensity-dependent effect can be removed by excluding low intensity spots from analysis, and/or

can be corrected using local weighted scatterplot smoothing (LOWESS) (Yang et al., 2002b). LOWESS also corrects for other systematic spatial variations, such as spotting inconsistencies and uneven hybridisation. There are other methods that apply local corrections, for example, local variance regularisation normalises the variance of the data (Yang et al., 2002a). Expression data are also commonly re-scaled by mean- or median-centring, so that the average $\log_2$ Cy5/Cy3 ratio of each gene across arrays is set to zero. This has the effect of enhancing the relative differences in expression ratios so that subsequent analysis will reveal the patterns rather than the abundance of gene expression (Quackenbush, 2001), as well as eliminating the effects of reference RNA. Specialist softwares are now available for microarray data processing. Some of these even allow users to perform a series of data filtering and normalisation methods in one package (Fielden et al., 2002; Saeed et al., 2003), which have greatly facilitated the processing of microarray data.

1.2.3.3 Identification of differentially expressed genes

One common application of microarrays is to identify differentially expressed genes between two samples, such as drug treated versus untreated cells (Wilson et al., 1999), or normal versus cancer tissue (DeRisi et al., 1996), where a binary system is a simple and efficient method. The two samples may be labelled with different fluorescent dyes and hybridised onto the same array, and the resultant expression ratios represent the changes in gene expression. Genes with $\log_2$ ratios closest to zero have relatively unchanged expression, whereas genes with very highly positive or negative $\log_2$ ratios are more abundant in one sample or the other. Such direct comparison requires dye swap hybridisations to eliminate the dye-bias effect. Alternatively, a common reference approach as mentioned above (section 1.2.2.2) can be used, and the $\log_2$ ratios of one sample compared to another. In either case, replicate arrays should be performed to provide an estimate for the variance in the data (Yang and Speed, 2002), and to enable identification of statistically significant up- or down-regulated genes. Various statistical tests have been used for microarray data, such as analysis of variance (ANOVA) (Churchill, 2004; Kerr et al., 2000), Wilcoxon rank sum test (Chambers et al., 1999). The Significance Analysis of Microarrays (SAM) method, a *t*-test based approach to assign scores for all genes and to calculate the false discovery rate (Tusher et al., 2001),

have been widely used to identify differentially expressed genes (Jones and Arvin, 2003; Larsson et al., 2005).

1.2.3.4 Expression patterns recognition by clustering analysis

A widely-used approach to study array data is by clustering analysis. The arrangement of genes into clusters according to the similarity of their expression patterns is very useful because genes that are co-regulated (similar expression patterns) are often involved in similar biological processes (Eisen et al., 1998). The power of the technique is maximised by clustering the data from multiple arrays so that samples with similar expression patterns are also grouped together (Tsai et al., 2005). This allows for identification of co-regulated genes across a series of conditions, common or contrasting patterns in different tissue types, or the gradual change of expression at increasing times post-treatment (section 1.2.4.3). Clustering analysis normally consists of 3 steps. The first step is the calculation of the pairwise distance between the expression vectors of the array elements, which are essentially composed of the Cy5/Cy3 ratios of the gene spots. Secondly, clustering algorithms group the elements into clusters by means of their pairwise distance. Thirdly, a visual representation of the clusters is created, usually an expression matrix in which genes are arranged in rows and arrays in columns as a dendrogram, similar to a phylogentic tree, in which the branch lengths represent the distances between the array elements. Various distance metrics exist for the calculation of the distances between array elements. For example, Euclidean distance is a method based on the Pythagorean Theorem. The Manhatten distance is the sum of distances between components of each expression vector. Pearson correlation, a semi-metric distance, may be centred, uncentred or squared (Quackenbush, 2001).

There exist many different clustering algorithms, of which the most popular are hierarchical clustering (Eisen et al., 1998), k-means clustering 1999 (Soukas et al., 2000; Tavazoie et al., 1999), and self-organising maps (Tamayo et al., 1999). Different algorithms have different characteristics, and when used on the same data set will result in slightly different clustering patterns. However, none of these algorithms is more 'correct' than the others; they simply reveal different features of the data set, hence

suited for different applications (Quackenbush, 2001). Hierarchical clustering (HCL) is an agglomerative approach, which gradually builds up a cluster by joining elements together based on their similarity. It begins by finding the two array elements with the smallest pairwise distance and joining them together. The distances between this newly joined group and each of the other elements or groups are calculated. The algorithm then repeats the 'search-and-group' process until all array elements are included in the cluster, forming a dendrogram (Quackenbush, 2001). Eisen *et al* first developed the programs Cluster and Treeview for the application of hierarchical clustering to microarray data analysis (Eisen et al., 1998), which became the standard analysis protocol. HCL has been used in analysing gene expression data of diseases such as psoriasis (Itoh et al., 2005), ischemic heart disease (Satterthwaite et al., 2005), heart failure (Steenman et al., 2005), primary biliary cirrhosis in liver (Honda et al., 2005), Parkinson disease (Hauser et al., 2005); in studying various cancers including HPV-associated cervical cancer (Vazquez-Ortiz et al., 2005), squamous cell carcinoma in lungs (Talbot et al., 2005). It has also been applied to analyse transcriptional profiling data of normal human tissues and organs (Shyamsundar et al., 2005), and microorganisms such as yeast (Eisen et al., 1998) and *Streptococcus pneumoniae* (Pandya et al., 2005). The advantages of HCL include the presention of results as a dendrogram, and that it does not require any prediction of the resulting clustering pattern. However, because HCL does not offer the option of pre-defined structure to the user, it has been alleged that biological interpretation of the HCL results is a complicated and laborious process (Okada et al., 2005).

In *k*-means clustering, the array elements are first randomly assigned into *k* clusters, a user-defined value (Tavazoie et al., 1999). The algorithm then calculates the distances within each cluster and between clusters. Then the array elements are shuffled between clusters in an iterative process, calculating the new distances after each move, until the array element arrives in a cluster where the distance is smallest. The process reiterates until no more movements can be made (Quackenbush, 2001). This algorithm is particularly suited for analyses where a partial structure is desired, as it allows for a pre-set number of clusters. Examples of *k*-means clustering applications include gene expression analyses to study human heart failure (Tan et al., 2002), stress-response in rats (Morgan et al., 2005), and the post-feeding processes in the malaria mosquito *Anopheles gambiae* (Dana et al., 2005).

Self-organising maps (SOMs) also allows the user to dictate the number of sections, in this case a simple geometric structure, such as a 2-dimensional rectangular grid. The nodes of the grid are 'trained' by an iterative process, adjusting their values so that they move closer towards their assigned genes, until the most effective separation is achieved (Quackenbush, 2001). The SOM algorithm has been incorporated into the program GENECLUSTER, (Tamayo et al., 1999). It has been used to analyse expression data in yeast (Tamayo et al., 1999) and *Clostridium acetobutylicum* (Alsaker and Papoutsakis, 2005). It has helped to identify gene expression patterns of the cellular response to HIV (Vahey et al., 2002), and to an oncogene, v-fos (Ordway et al., 2005). It has also been applied to studying the development and molecular signature of many cancer cells, including squamous cell cancer (Hummerich et al., 2005), Hepatitis C virus-associated hepatocellular carcinoma (Iizuka et al., 2005; Tsunedomi et al., 2005), promyelocytic leukaemia (Song et al., 2003), B-cell lymphomas (Wang et al., 2002b), embryonal carcinoma (Harris and Childs, 2002), as well as the molecular classification of tumour type (Covell et al., 2003).

The Institute for Genome Research (TIGR) have developed a suite of programs (TM4) for microarray analysis, which includes an application for data processing and normalization, Microarray Data Analysis System (MIDAS); also a program for the analysis and visualisation of microarray data, Multiexperiment Viewer (MeV). Both applications have incorporated a large number of algorithms, allowing users to process, normalise, cluster and perform statistical analysis for microarray data.

1.2.3.5 Confirmation of microarray results

Like many other scientific experiments, it is important that microarray studies include the use of additional methods to verify the results. Usually a small subset of the genes identified by microarrays are confirmed using follow-up methods, for example reverse transcription PCR (RT-PCR) (Corbeil et al., 2001; Ryo et al., 2000; Tang et al., 2005; van 't Wout et al., 2003), quantitative real-time RT-PCR (Carter et al., 2002; Leong et al., 2005), and northern blotting (Ahn et al., 2002; Chambers et al., 1999; Chang and Laimins, 2000; Paulose-Murphy et al., 2001; Ryo et al., 2000). The results from these

methods are gnerally consistent with the microarray results, hence strengthening the reliability of the microarray findings. Methods such as RT-PCR may also differentiate expression patterns of splice variants (Menegazzi et al., 1999; Mirandola et al., 1998). Moreover, since changes in transcript abundance may not necessarily result in changes in protein levels, in some experiments western blotting (Guerra et al., 2004) and immunoassays for proteins (Tang et al., 2005) have been used to correlate transcriptional changes identified by microarrays with protein levels.

1.2.4 Applications of microarrays in host-pathogen interactions

During all stages of infection, viruses subvert the host cellular machineries for their own benefit, altering the expression of many cellular genes. Meanwhile host organisms have developed many ways to control the infection. At any given time, the expression patterns of both viral and cellular genes represent the complex interactions between host and pathogen. The use of microarray techniques in studying these host-pathogen interactions during infection has drawn particular interests (Cummings and Relman, 2000; Diehn and Relman, 2001; Jenner and Young, 2005; Kellam, 2001), and has provided many insights in host-pathogen studies.

1.2.4.1 Microarray studies of microbial gene expression

The availability of microbial genome sequences (section 1.2.1) have prompted the development of whole- (or almost whole) genome microarrays for many microorganisms, and the gene expression of *Escherichia coli* (Richmond et al., 1999), *Mycobacterium tuberculosis* (Wilson et al., 1999), *Plasmodium falciparum* (Hayward, 2000; Rathod et al., 2002), *Streptococcus pneumoniae* (de Saizieu et al., 2000), *Salmonella typhimurium* (Eckmann et al., 2000; Rosenberger et al., 2000), and *Entamoeba histolytica* (MacFarlane et al., 2005b) have been studied. The arrays used to study the physiological state of *E. coli* during heat shock discovered expression changes in more than 80 previously known heat-shock related-genes and 35 novel genes (Richmond et al., 1999). The temporal transcription programme during the metabolic shift from anaerobic fermentation to aerobic respiration in *Saccharomyces cerevisiae* has also been investigated using microarrays (DeRisi et al., 1997). Microarrays also revealed coordinated gene expression in *Plasmodium falciparum*, the malarial parasite (Ben Mamoun et al., 2001). Genes involved in protein synthesis were expressed during the early stages of infection, followed by metabolism genes, and genes involved in adhesion and invasion. Microarrays are also useful in studying drug actions. A study in *M. tuberculosis* discovered that four genes encoding fatty acid biosynthesis enzymes were co-induced with a known target of isoniazid in the presence of the drug. Therefore these genes may be potential targets for new anti-tuberculosis drugs (Wilson et al.,

1999). Oligonucleotide arrays have been used to detect drug-resistant mutants of *M. tuberculosis* (Troesch et al., 1999). Based on the observation that genes specifically induced during infection are often required for virulence (Cotter and Miller, 1998), probable virulence factors can be identified by comparing gene expression profiles of virulent and avirulent strains. For example, microarrays have found that virulent isolates of *Vibrio cholerae* from patient stool had upregulated expression of 44 genes compared to non-pathogenic laboratory strains (Merrell et al., 2002). Many of these genes were involved in cell motility and nutrient acquisition.

1.2.4.2 Microarray studies of viral gene expression

Microarray applications are particularly suited for analysing viruses with large genomes, for the same reasons as other microbes (section 1.2.4.1). Both spotted DNA microarrays and oligonucleotide microarrays have been put to use in studying viral infections. The transcriptomes of many herpesviruses, including herpes simplex virus (HSV) (Stingley et al., 2000), HCMV (Chambers et al., 1999), HHV-6 (Ohyashiki et al., 2005), Kaposi-sarcoma associated herpesvirus (KSHV) (Jenner et al., 2001; Paulose-Murphy et al., 2001) and murine gammaherpesvirus 68 (MHV-68) (Ahn et al., 2002), have been analysed by microarrays. These microarray studies revealed that herpesvirus genes were transcribed in a co-ordinated manner, with viral genes involved in particular functions co-expressed at the corresponding stages of infection. Most of these array studies have confirmed the kinetic classification of herpesvirus genes and suggested putative functions of unclassified viral genes. For example, Paulose-Murphy *et al* predicted that 4 KSHV ORFs (11, 31, 35, 49) may function as transactivators, one (ORF10) may be involved in viral DNA replication, and 12 ORFs (K7, 18, 23, 24, 27, 28, 30, 33, 42) were possible structural genes (Paulose-Murphy et al., 2001). In another study (Jenner et al., 2001), a region between K10.5 and K11 in KSHV was identified by microarrays, which may represent a novel transcribed ORF (K10.7) with homology to viral interferon regulatory factors (vIRFs). More details on temporal classification of herpesvirus using microarrays are given in section 5.1. Microrrays have also been produced for VZV (Kennedy et al., 2005) and vaccinia virus (Wenli et al., 2004), which are sensitive to viral transcripts, although these have not yet been used in expression profiling.

1.2.4.3 Microarray studies of host gene expression during viral infection

The expression patterns of cellular genes during infection by various viruses have been explored using microarrays (Table 1.4). Amongst these most followed the infection *in vitro*, except a few studies on *in vivo* (Bigger et al., 2001; Domachowske et al., 2002; Taylor et al., 2000). Data showed that viral infection had substantial influence on many cellular processes, not only on defence mechanisms such as DNA repair, cytokine signalling and stress response, but also on normal cellular functions like transcription regulation, mRNA processing, protein synthesis, protein degradation, cell cycle progression and mitochondrial functions (Jenner and Young, 2005). Very often hundreds of genes are significantly upregulated or downregulated during infection (Corbeil et al., 2001; Zhu et al., 1998), demonstrating the complexity of viral infection, also the power of microarray application in host-pathogen interaction analysis.

Microarray technology offers powerful comparative studies. For example, experiments have compared viruses capable of replication with their inactivated counterparts. These have helped to identify host responses that are dependent on viral replication (Browne et al., 2001; Geiss et al., 2001; Jones and Arvin, 2003; Mossman et al., 2001; Wang et al., 2005). In particular, it has been found that replicating HCMV induces fewer interferon-responsive genes and proinflammatory cytokines in human fibroblasts than their non-replicating counterparts (either by UV-inactivation or by treatment with metabolic inhibitor), suggesting that viral gene products may block cellular antiviral responses (Browne et al., 2001). Alternatively, more changes in cellular gene expression were caused by replicating influenza virus than by inactivated virus (Geiss et al., 2001). Studies have also compared the cellular responses to related viruses that cause different clinical symptoms, offering insights into the pathogensis of the more virulent virus. For example, it was found that the severe acute respiratory syndrome-associated coronavirus (SARS-CoV) induces greater changes in the transcription of cellular genes associated with apoptosis, inflammation, stress response and procoagulation than human coronavirus 229E which causes the common cold (Tang et al., 2005). Differential gene expression has been observed between the oncogenic adenovirus type 12 and non-oncogenic adenovirus type 5 (Vertegaal et al., 2000), also between patient samples

obtained from hepatitis B and C lesions (Honda et al., 2001). A study has also compared the gene expression patterns in different cell types to identify cell type-specific responses to VZV (Jones and Arvin, 2003).

Some have studied how host gene expression were affected by knocking out viral genes, such as HSV-1 UL54 (Stingley et al., 2000) and HCMV U28 (Hertel and Mocarski, 2004); and by expressing specific virus proteins such as HIV Tat (de la Fuente et al., 2002), HIV Nef (Shaheduzzaman et al., 2002; Simmons et al., 2001), KSHV LNA-1 (Renne et al., 2001), KSHV vGPCR (Polson et al., 2002), and IE86 of HCMV (Song and Stinski, 2002). Such studies have helped identify host responses to specific viral proteins, for example HCMV glycoprotein B been shown to trigger IFN-stimulated genes (Simmen et al., 2001).

In summary, microarrays are a powerful tool for the large-scale monitoring of changes in gene expression that have been extensively used in many applications, and have provided many useful insights in cell biology, disease, and host-pathogen interactions.

Table 1.4 Microarray studies of host gene expression in viral infections.

Family	Virus species [a]	Cell type [b]	Array type [c]	No. probes	Cellular response category [d]							Reference
					1	2	3	4	5	6	7	
dsDNA												
Adenoviridae	AdV type-5 and type-12	HEL, HEK, HER cells	M	18376			+					Vertegaal et al 2000
Herpesviridae	EBV	B-lymphoma cell lines	M	4146	+		+/-		+/-	+		Carter et al 2002
	HCMV	HFF	O	12626	+	+	+/-		-	+		Browne et al 2001
	HCMV	HFF	S	8942	+							Simmen et al 2001
	HCMV	HFF	O	1650x4	+	+		+	-	+	+/-	Zhu et al 1998
	HCMV	HFF	S	29593	+			-		+	mitosis	
	HHV6	SupT1	M	1152		+/-						Mayne et al 2001
	HSV-1	HEL fibroblasts	M	588							arrest G2-M	Hobbs and DeLuca 1999
	HSV-1		M	588				+	-		arrest G2-M	Khodarev et al 1999
	HSV-1	HEL fibroblasts	S	19000	+			+				Mossman et al 2001
	HSV-1		S	57	+						+/-	Stingley et al 2000
	HSV-1		M	18378						+	+/-	Tsavachidou et al 2001
	VZV	T-cells, fibroblasts, human skin grafts in mice	S	9180	+		+/-	+/-	+/-		+/-	Jones and Arvin 2003
	KSHV	DMVEC	S	4165	+			+			mitosis	Moses et al 2002
	KSHV	DMVEC	M	2350	+			+				Poole et al 2002
	Marek's disease virus	Chicken embryo fibroblasts	M	1126	+							Morgan et al 2001
Poxviridae	Vaccinia virus	HeLa	S	15360	-			-		-		Guerra et al 2003
	Vaccinia virus Ankara	HeLa	S	15360	+	+						Guerra et al 2004
Papovaviridae	HPV 31	NHK	S	7075	-							Chang and Laimins 2000
dsDNA-r												
Hepadnaviridae	Hepatitis B virus	Liver cells from patients	S	1080	+	+	+	+				Honda et al 2001
	Hepatitis B virus	HCCs from patients	S	23040							proliferation	Okabe et al 2001
dsRNA												
Reoviridae	Reovirus serotypes 1 & 3	HEK293	O	12626							arrest G2-M	Poggioli et al 2002
	Rhesus rotavirus	Caco3	S	38432	+	+		+	+			Cuadras et al 2002
ssRNA-												
Orthomyoxoviridae	Influenza virus	HeLa	S	4608		-		+/-	+/-	-		Geiss et al 2001
Paramyoxviridae	Measles virus	PBMCs	M	1176x3	+			+				Bolt et al 2002

Family	Virus species [a]	Cell type [b]	Array type [c]	No. probes	Cellular response category [d] 1	2	3	4	5	6	7	Reference
	Pneumonia virus of mice	Lung tissues from mice	O	12000	+	+		+				Domachowske et al 2002
Paramyoxviridae	Respiratory syncytial virus	A549	M	268		+		+		+		Zhang et al 2001b
	Respiratory syncytial virus	Hep2	O	12626	+			+	-			Tian et al 2002
	Respiratory syncytial virus	Hep2, A549	O	12626	+							Zhang et al 2001b
ssRNA+												
Flaviviridae	Hepatitis C virus		O	7000	+							Bigger et al 2001
	Hepatitis C virus	Liver cells from patients	S	1080	-	-					+/-	Honda et al 2001
	Hepatitis C virus	HCC from patients	S	23040							Anti-cancer	Okabe et al 2001
Picornaviridae	Coxsackievirus b3	Heart tissues from mice	S	7000				+				Taylor et al 2000
	Echovirus 1		M	588				+			arrest G2-M	Pietiainen et al 2000
	Poliovirus	HeLa	S	10000				+				Johannes et al 1999
Coronaviridae	SARS-CoV	Vero E6	S	16000	+/-		+/-	+/-		+/-	+/-	Leong et al 2005
	SARS-CoV	Huh7	O	22000	+	+	+/-	+		+/-	+/-	Tang et al 2005
	Coronavirus 229E	Huh7	O	22000				+/-		-	+/-	Tang et al 2005
SsRNA-RT												
Retroviridae	HIV-1	human T-cell line CEM-CCRF	S	1506			+/-					Geiss et al 2000
	HIV-1		O	6800	+			+	-		arrest G2-M	Corbeil et al 2001
	HIV-1	human T-lymphoblastoid cell lines	S	4608	+/-	+/-	+/-	+/-	-	-	arrest G2-M	van't Wout et al 2003
	Human foamy virus	HEL	M	588							arrest G2-M	Wagner et al 2000
	Human T-cell leukaemia virus		M	588				+		+	progression	de La Fuente et al 2000

[a] Virus species: Adv, adenovirus; EBV, Epstein-Barr Virus; HCMV, human cytomegalovirus; HHV-6; human herpesvirus type-6; HSV-1, herpes simplex virus 1; VZV, varicella-zoster virus; KSHV, Kaposi's sarcoma-associated herpesvirus; HPV, human papillomavirus; SARS-CoV, severe acute respiratory syndrome-associated coronavirus; HIV-1, human immunodeficiency virus-1

[b] Cell type: HEL, Human embryonic lung fibroblasts; HEK, human embryonic kidney fibroblasts; HER, human embryonic retina fibroblasts; HFF, Human foreskin fibroblasts; SupT1, Human T-cell line; DMVEC, Dermal microvascular endothelial cells; NHK, Normal human keratinocytes; HeLa, ; PBMCs, peripheral blood monocytic cells; Huh7, Human hepatoma cell line; CEM-CCRF, Human T-cell line; Hep2, ; A549, Human lower airway epithelial cells; HCC, hepatocellular carcinoma; Caco3, Human intestine cell line, Vero E6, African green monkey kidney cells.

[c] Array type: M, membrane array; S, spotted DNA microarray; O, Oligonucleotide chip (synthesised *in situ*).

[d] Cellular response category: 1, Interferon and pro-inflammatory response; 2, Cytokines; 3, Signal transduction; 4, Stress response (apoptosis, heat shock, oxidative stress); 5, Protein synthesis; 6, Transcription factors; 7, Cell cycle. +, expression mostly upregulated; -, expression mostly downregulated; +/-, expression altered

1.3 Aim and Scope of thesis

This PhD thesis aims to investigate the complex, intertwined processes during lytic infection of the human T cell by human herpesvirus type-7 (HHV-7), a highly prevalent and relatively non-pathogenic virus. I aim to study the expression of viral and host genes simultaneously, in order to identify the cellular responses to HHV-7 at different stages of lytic infection. I also intend to examine the lytic replication cycle of HHV-7 and compare this to other human betaherpesviruses. To achieve this, the following steps were undertaken:

1) A custom-made host-pathogen microarray was created in-house, and its proficiency for use in measuring HHV-7 and host gene transcription was tested (Chapter 3).
2) An *in vitro* system for cell-free HHV-7 infection of human T-cells with a high multiplicity of infection was developed to allow the study of the HHV-7 gene expression programme (Chapter 4).
3) Finally, the custom made microarrays were applied to the analysis of HHV-7 and human cellular gene expression in parallel. Particularly, the temporal regulation of HHV-7 viral gene expression was examined by expression profiling with and without the use of metabolic inhibitors on *de novo* protein synthesis and viral DNA replication (Chapter 5).

So far no extensive gene expression studies on HHV-7 have been performed, as such studies are limited by the virus's low infectivity. This is the first description of a host-pathogen microarray analysis for HHV-7.

CHAPTER 2

MATERIALS & METHODS

2.1 Buffers and Solutions

Table 2.1 Constituents of buffers and solutions

Solution	Ingredients
Deoxynucleotide triphosphate mix (dNTPs)	100 mM of deoxyadenosine triphosphate (dATP), deoxythymidine triphosphate (dTTP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP)
Luria-Bertani (LB) agar	1% (w/v) bacto tryptone, 0.5% (w/v) bacto yeast, 0.5% (w/v) sodium chloride, pH 7.0 with 1.5 % (w/v) bacto-agar
Luria-Bertani (LB) broth	1% (w/v) bacto tryptone, 0.5% (w/v) bacto yeast, 0.5% (w/v) sodium chloride, pH 7.0
10x PCR buffer (20 mM magnesium chloride)	100 mM Tris, pH 8.3, 500 mM potassium chloride, 20 mM magnesium chloride
10x PCR buffer (10 mM magnesium chloride)	100 mM Tris, pH 8.3, 500 mM potassium chloride, 10 mM magnesium chloride
Phosphate-buffered saline (PBS)	137 mM sodium chloride, 2 mM potassium chloride, 10 mM sodium hydrogen phosphate (dibasic), 2 mM potassium hydrogen phosphate (dibasic), pH 7.4
20x Saline sodium citrate (SSC)	3 M sodium chloride, 0.3 M sodium citrate, pH 7.0
Saline-sodium phosphate-EDTA buffer SSPE	2.98 M sodium chloride, 0.02 M EDTA, 0.2 M phosphate buffer, pH 7.4
Tris-acetate-EDTA (TAE)	40 mM Tris pH 7.8, 20 mM sodium acetate, 1 mM ethylenediaminetetraacetic acid (EDTA)
Tris-EDTA (TE)	10 mM Tris pH 7.4, 1 mM EDTA

2.2 Mammalian cell culture

2.2.1 Cell lines

Table 2.2 Cell lines

Cell line	Description
CEM-SS	Human T4-lymphblastoid cell line, suspension
HeLa	Cervical epithelium, adherent
Huh7	Hepatoma, adherent
MRC5	Embryonic fibroblasts, adherent
Ramos	B-cell, suspension
SupT1	Non-Hodgkin's T-cell Lymphoma, CD4+, suspension
U937	Monocytes, suspension

2.2.2 Cell growth media

Suspension cell lines were grown in Roswell Park Memorial Institute (RPMI)-1640 medium (Invitrogen, UK) supplemented with 10% foetal bovine serum (FBS, Biowest). When infected with virus, cultures were maintained in RPMI with reduced (5%) FCS to slow down cell growth. Adherent cell lines were cultured in Dulbeccos Modified Eagle

Media (DMEM) medium (Invitrogen, UK) with 5% FCS. All culture media were supplemented with 100 units/ml penicillin and 100 μg/ml streptomycin (P/S, Invitrogen, UK). All mammalian cells were cultured in tissue culture flasks with filtered caps or multi-well plates (Helena Biosciences, UK) in humidified incubators, 37 °C, 5% CO_2. Centrifugations of cells were carried out at 325g, 5 min, 20 °C, unless otherwise stated. All cells were tested negative for mycoplasma.

2.2.3 Thawing

Cells taken from liquid nitrogen were thawed immediately at 37 °C and added to 2 × 5 ml pre-warmed medium. Within 24 h cells were washed and medium replaced. Suspension cells were pelleted at 325g, 5 min and washed in 10 ml medium, followed by centrifugation for 5 min and resuspension in pre-warmed medium at 2×10^5 cells/ml. For adherent cells medium was removed and cells were washed with 10 ml medium per flask, and fresh medium was added to 1×10^5 cells/ml.

2.2.4 Passaging cells

Cell counts were obtained using a haemocytometer. Suspension cells were split 1:4 to 1:8, every 3-6 days, depending on cell line and cell density. To split adherent cell lines, cells were first detached from culture surface using 1 ml trypsin/EDTA (Invitrogen, UK) per 25 cm^2 culture area, and then split 1:2 to 1:6.

2.2.5 Freezing cells

Cells were pelleted by centrifugation and resuspended in corresponding freezing medium (FCS with 10% dimethyl sulphoxide (DMSO, Sigma, UK) at a final concentration of $1\text{-}2\times10^6$ cells/ml. Cell stocks were aliquoted into cryovials (Nunc, USA) and frozen gently to -80 °C overnight in an isopropanol-containing cryo-container (Nalgene, USA) before being transferred to liquid nitrogen for storage at -140 °C.

2.3 Culturing HHV-7

2.3.1 Virus strain

Human herpesvirus type 7 (HHV-7) used in this thesis was strain DC, provided by Dr Duncan Clark, Royal Free & University College Medical School, London.

2.3.2 Virus infection

For infection of SupT1 cell line by cell-free human herpesvirus type 7 (HHV-7), 1 ml of virus stock was allowed to adsorb to 1×10^5 – 5×10^6 cells for 1-2 h at 37 °C. For spinoculation, the virus-cell mixture was centrifuged at 1000 g for 1 h at 20 °C. After adsorbtion (with or without centrifugation), the virus-cell mixture was seeded into tissue culture flasks with pre-warmed medium and incubated at 37 °C, 5% CO_2.

2.3.3 Virus culture and maintenance

Virus culture was maintained by co-culturing as follows. HHV-7 infected cells were monitored for the display of cytopathic effect (CPE), characterised by enlarged, multinucleated cells, and elongated cells attached to bottom. Cultures with greater than 50% CPE were co-cultured with uninfected cells at the ratio of 1:1 to 1:5 (infected: uninfected) by cell count. Medium was added to achieve cell density of 2×10^5 – 1×10^6 cells/ ml.

2.3.4 Stock preparation

HHV-7 virus stocks were prepared as follows. HHV-7 infected cells were centrifuged and resuspended in 1/6 – 1/100 original volume of FBS. Cells were lysed by freeze-thawing 6 times to release cell-associated virus particles. Cell lysates were centrifuged

again to separate cell debris. Supernatant containing virus particles was retrieved, aliquoted into 1ml stocks, and stored at –140 °C (liquid nitrogen).

2.3.5 Virus quantification

2.3.5.1 Observation of cytopathic effect (CPE)

HHV-7 infected SupT1 cells were monitored for cytopathic effect (CPE) by visual observation under light microscopy (200× magnification). Percentage CPE was estimated from the proportion of cells attached to bottom of tissue culture flask, enlarged (2× normal diameter or above), elongated or become irregularly shaped (detailed in section 4.2.2).

2.3.5.2 Immunofluorescence assay (IFA)

HHV-7 infected cells were harvested and centrifuged, washed in 2× original volume of Phosphate buffered saline (PBS) and resuspended in PBS to a final concentration of 1×10^6 cells/ml. 2×10^4 cells were then spotted onto each well of polytetrafluoroethylene (PTFE)-coated multispot slides (C.A. Hendley Ltd), air-dried for 40 min, and fixed in cold acetone (-20 °C) for 10 min. Fixed cells were then assayed for HHV-7 infection by incubating with mouse-derived anti-HHV-7 monoclonal antibody that target an early HHV-7 antigen (20 µl per well at 2.5 ng/µl in PBS; H2034-18, US Biological, Massachusetts) for 40 min in humidified chamber at 37 °C. The slide was then washed three times in PBS, followed by incubation with FITC-conjugated anti-mouse IgG antibody (20 µl per well at 5 ng/µl in PBS; F2883, Sigma) for 40 min, 37 °C. Slides were washed 3 times in PBS and air-dried. Slides were then counterstained with Vectashield mounting medium with DAPI (Vector Laboratories, USA) and examined by fluorescent microscopy using the Axioplan2 imaging Pol (Zeiss, Germany), to detect the presence of HHV-7 antigen (green fluorescent). The number of green fluorescent (by FITC) cells and the total number of cells (by DAPI counter-stain) were counted in four fields per well. The mean percentage of infected cells was calculated as:

$$\% \text{ Infected cells} = \frac{\text{HHV-7 antigen-positive cells (FITC)}}{\text{Total cells (DAPI counter stain)}}$$

2.3.5.3 Infectious units count (IUC)

SupT1 cells were plated in a 24-well plate at 5×10^5 cells per well and pelleted by centrifugation. 100 µl of serially diluted virus (neat, ½, ¼) were added per well in triplicate per dilution and incubated at 37 °C for 2 h to allow virus attachment. 500 µl of medium (RPMI, 5% FBS, P/S) was then added to each well and the cells were incubated at 37 °C for 3-4 days. Cells were then observed for CPE (section 2.3.5.1) or harvested for IFA (section 2.3.5.2). Infectious units count per ml virus (by CPE or by IFA) was calculated as:

$$\% \text{ Infected cells} \times 5\times10^5 \text{ cells} \times \frac{1000\ \mu l}{100\ \mu l}$$

2.3.5.4 $TCID_{50}$ determination

SupT1 cells were plated in 24-well plates at $1\times10^5 - 5\times10^5$ cells per well and pelleted by centrifugation. 100 µl of serially diluted virus (half-log dilution $10^0 - 10^{-5.5}$) were added per well in quadruplicate per dilution and incubated at 37 °C for 1-2 h to allow virus attachment. 100 µl of RPMI-10 medium was then added to each well and the cells were incubated at 37 °C for 3-4 days. After incubation cells were observed for CPE (section 2.3.5.1) or harvested for immunofluorescence detection of infection (section 2.3.5.2). Wells containing unambiguously infected cells were scored as positive and $TCID_{50}$ was calculated as follows:

$$-\log TCID_{50} = -\log L - \log d(S - 0.5)$$

Where: L = highest virus dilution to give 100% positive

d = dilution factor (for half-log dilution, $d = 0.5$)

S = sum of the ratios (positive wells/total wells) for every dilution from L

2.3.6 Flow cytometry

Cells were prepared for flow cytometry analysis as follows. 5×10^5 cells were washed by adding 2× volume PBS and spinning at 1300 g, 5 min, 4 °C. Cells were fixed on ice for 10 min in 4% paraformaldehyde in PBS. Cells were washed twice in PBS and permeabilised in FACS permeabilisation solution (0.1% Triton X-100, 0.01% sodium azide, 2% FBS in PBS). Permeabilised cells were incubated with anti-HHV-7 monoclonal antibody (2.5 ng/µl; H2034-18, US Biological, Massachusetts) in 50 µl FACS perm + 10% FBS for 1 – 2 h on ice. Unbound antibody was removed by washing three times in 1 ml FACS permeabilisation solution. Cells were incubated with the secondary antibody, FITC-conjugated anti-mouse IgG (5 ng/µl; F2883, Sigma) in 50 µl FACS perm + 10% FBS for 1 – 2 h on ice in the dark. Finally cells were washed three times in 1 ml FACS permeabilisation solution and once in 1 ml FACS wash solution (0.01% sodium azide, 2% FBS in PBS). Ten thousand events of the prepared cells were collected and analysed using a FACScan with Cellquest software (Becton Dickinson, UK).

2.3.7 Time course infection of SupT1 cells by HHV-7

For each time point, 1×10^5 cells of SupT1 (per well in 24-well plate) were inoculated with 1 ml HHV-7 virus stock (MOI = 2.5) or mock infected with uninfected cell lysate for 1 h, with spinoculation (1000 g) at 20 °C. After 1 h the virus inoculum was carefully removed without disturbing the cells. 1 ml of RPMI medium (P/S, 5% FBS) was added to the cells and the plates incubated at 37 °C. At each time point (8, 12, 16, 20, 24, 38, 48, 72 hours post infection) the cells (HHV-7-infected and mock-infected) were centrifuged to remove supernatant. The cells were resuspended in 350 µl RLT lysis buffer (Qiagen, UK) and stored at -80 °C for later RNA extraction (section 2.7.2). In earlier attempts when 5×10^6 cells were used per time point, cells were lysed in 1ml TRIzol reagent (Invitrogen, UK) and stored at -80 °C for later RNA extraction (section 2.7.1).

2.3.8 Metabolic inhibitor assay for temporal gene expression mapping

For metabolic inhibitor assay, SupT1 cells were first treated with 200 μg/ml cycloheximide (CHX, Sigma, UK), 500 μg/ml phosphonoacetic acid (PAA, Sigma, UK), or mock (medium) for 1 h at 37 °C prior to infection by HHV-7. Cells were then centrifuged to remove supernatant, and inoculated with HHV-7 stock as in section 2.3.7. Constant concentration of inhibitors was maintained throughout the infection and time course. Cells were harvested at 8 and 16 hpi (CHX-treated and untreated); at 24 and 48 hpi (PAA-treated and untreated) in 350 μl RLT lysis buffer (Qiagen, UK) and stored at -80 °C for later RNA extraction (section 2.7.2).

2.4 Producing HHV-7 array probes

2.4.1 Array probe primers

A list of primers used in this thesis is provided in Table 3.1.

2.4.2 HHV-7 DNA preparation

To obtain template DNA for amplification of HHV-7 probes, DNA was isolated from HHV-7 infected SupT1 cells using the Wizard Genomic DNA Purification Kit (Promega, UK). Briefly, 2×10^6 SupT1 cells (2 ml, infected with HHV-7 with CPE> 50%) were harvested by centrifugation and washed with 1 ml PBS. Cells were then centrifuged at 13,000 g to remove the supernatant, and the cell pellet was resuspended in 200 µl PBS. Cells were lysed with 600µl of nuclei lysis solution. RNA was digested using 3µl RNase solution for 20 min at 37°C, and protein was removed by incubating with 200µl protein precipitation solution on ice for 5 min. The precipitated protein was removed by centrifugation at 16,000 g, 20°C for 4 min. The DNA in the supernatant was recovered into a clean tube which contained 600µl isopropanol (20°C) and was centrifuged at 16,000 g, 20°C for 1 min to precipitate the DNA. Supernatant was removed and the DNA pellet washed with 600µl 70% ethanol (20°C), centrifuged again to remove the ethanol and air-dried for 10 min. The DNA pellet was resuspended in 100µl of DNA rehydration solution (10mM Tris-HCL (pH 7.4), 1mM EDTA (pH8.0)) at 4°C overnight. The concentration of the purified DNA was calculated from the UV absorbance at 260 nm using a UV spectrophotometer (Camlab, UK). An absorbance of 1 cm^{-1} was taken to be equivalent to 50 µg/ml DNA. DNA was stored at -20°C until use in PCR (section 2.4.3).

2.4.3 Initial amplification of viral DNA

Polymerase chain reaction (PCR) was performed to generate array probes of approximately 300 bp in length for each open reading frame (ORF) of HHV-7 genome.

Probes complimentary to the 5'-end of each ORF (Figure 3.2) were designed using the Primer3 program, so that all probes have similar length and annealing temperature. Each PCR reaction contained 500 pmol of each primer (Table 2.1, Oswel Research Products Ltd., UK), 1.5 μg of HHV-7 DNA (section 2.4.2), 20 μmol dNTP (Promega), and 1.25 units Tag DNA polymerase (Promega), in 50 μl reaction buffer (10 mM Tris (pH8.3), 50 mM KCl, 2 mM $MgCl_2$ in UV-sterilised water). PCR reactions were carried out in Peltier thermal cycler (MJ Research, USA) with the following programme: 95 °C for 3 min; 30 cycles of 94 °C for 30 sec, 54 °C for 30 sec, 72 °C for 45 s; then a final extension step at 72 °C for 7 min. The presence and size of PCR products were checked by agarose gel electrophoresis (section 2.4.4) and if satisfactory, were used in ligation (section 2.4.5.2) directly.

2.4.4 Agarose electrophoresis of PCR product

PCR amplified probes were analysed by gel electrophoresis using 1.5 % agarose gel Electrophoresis analysis of nuclei acids were performed in 1-2% agarose gel in 1x Tris-acetate-ethylenediaminetetraacetic acid (EDTA) (TAE), containing 0.2 μg/ml ethidium bromide (Sigma, UK), under electrical potential difference of 60-120V (100-200 mA). Separated nucleic acids were visualised under UV illumination.

2.4.5 Cloning array probes

2.4.5.1 Competent cell preparation

Escherichia coli (XL1-Blue) (Stratagene, UK) was grown to saturation at 37 °C in a shaking incubator overnight in 5 ml Luria-Bertani (LB) broth containing tetracyclin (10 μg/ml). This was used to inoculate LB-broth (1:100) and shaken at 37°C until the culture reached mid-log phase (OD_{550} ~0.5, as determined by spectroscopy). The culture was chilled on ice for 10 min, followed by centrifugation at 3000 g, 5 min, 4 °C. The bacteria were resuspended in half original volume of $CaCl_2$ (50 mM CaCl2, 10 mM Tris pH8.0, 4 °C) and left on ice 15 min. Bacteria were then centrifuged again and resuspended in 1/15th original volume of $CaCl_2$. Competent bacteria were either used

immediately or stored frozen in liquid nitrogen in 200 µl aliquots (100 mM calcium chloride containing 15% (v/v) glycerol).

2.4.5.2 Ligation of DNA into pGEM-T easy vector

The HHV-7 PCR products (2.4.3) were ligated directly into 50 ng of pGEM-T Easy vector (Promega, UK), using 3 units of T4 DNA ligase (Promega, UK), 5 µl 2×ligase buffer (20mM Tris-HCl (pH 7.8), 10mM $MgCl_2$, 10mM DTT, 1mM ATP) in 10 µl at 4°C for 16 h.

2.4.5.3 Transformation

The ligation reaction was incubated with 200 µl of competent XL-1 Blue *E.coli* (section 2.4.5.1) on ice for 30 min. Uptake of DNA was induced by heat shock at 42 °C, 45 sec, after which cells were returned to ice immediately and cooled for 2 min. Transformed cells were then allowed to equilibrate in 1 ml LB without antibiotics at 37 °C, 30 min. Cells were then pelleted at 3000 g, 10 min, 20 °C, and resuspended cells in residual LB (approx. 200 µl) before spreading onto LB-agar plates containing 100 µg/ml ampicillin (Sigma, UK), 0.5 mM isopropylthio-β-D-galactoside (IPTG, Promega, UK), 80 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal, Promega, UK). The plates were incubated at 37 °C for 16 h to select for plasmid-containing, ampicillin-resistant bacteria, and for the insert-containing white colonies by X-gal/IPTG screening.

2.4.6 Identification of array probes

2.4.6.1 Growth of bacterial clones

To confirm the viral DNA inserts, 5 white colonies (indicate presence of transformed PCR product) were picked from each plate (section 2.4.5.3) and screened for the presence of viral DNA insert. Each colony picked was first dipped into PCR reaction mixture (section 2.4.6.1) and then into 150µl LB (100 µg/ml ampicillin) on 96-well

plate and allowed to grow overnight at 37 °C. Next day the clones that were confirmed positive (contain viral DNA inserts) by PCR were selected for DNA extraction (section 2.4.6.2) and stored as archive in 20% glycerol v/v in 96-well plate at -80 °C. A maximum of two aliquots per clone were stored.

2.4.6.2 Screening for viral insert-containing clones by PCR

Polymerase chain reaction was performed to confirm the presence of viral gene inserts in XL1-Blue transformants using universal pGEM-T forward and reverse primers (Table 2.1). The ingredient of each PCR reaction was the same as in section 2.4.3, except for primer (150 pmol of each pGEM-T primer, Thermal BioSciences GmbH, Germany) and template (1 white colony of transformed XL1-Blue). PCR reaction was carried out as follows: 95 °C for 5 min; 30 cycles of 94 °C for 30 sec, 55°C for 30 sec, 72 °C for 1 min; then 72 °C for 7 min. Products were checked for the presence of viral inserts by agarose gel electrophoresis (section 2.4.4).

2.4.6.3 Purifying plasmid DNA from positive clones

Plasmid DNA was isolated from the viral insert-containing *E.coli* (XL1-Blue) clones using the QIAprep Spin Miniprep kit (Qiagen, UK), a modified alkaline lysis method (Birnboim and Doly, 1979). 1.5 ml of bacterial culture (section 2.4.6.1) was pelleted at 14,000 g for 10 min and the bacteria resuspended in 250 µl resuspension buffer P1 (50 mM Tris-Cl (pH 8.0), 10 mM EDTA, 100 µl RNase A). Bacteria were lysed with the addition of 250 µl lysis buffer P2 (200 mM NaOH, 1% SDS (w/v)), and the solution neutralised with 350 µl neutralisation buffer N3 (3.0 M potassium acetate (pH 5.5). Genomic DNA was pelleted by centrifugation at 14,000 g for 10 min and the supernatant applied to a QIAprep column. The solution was passed through the column by centrifugation at 14,000 g, 1 min. The column was washed once with 500 µl buffer PB, twice with 750 µl buffer PE, and plasmid DNA eluted with 50 µl distilled water. The concentration of the purified DNA was calculated from the UV absorbance at 260 nm using a UV spectrophotometer (Camlab, UK). An absorbance of 1 cm^{-1} was taken to be equivalent to 50 µg/ml DNA.

2.4.6.4 Verifying sequence of viral insert

The identity of the clones were verified by DNA sequencing using an automated application of the chain-termination method (Sanger et al., 1977). 1 µl of plasmid DNA was added to 4 µl of Beckman QuickStart mix (containing dATP, dCTP, dTTP, dITP, ddUTP, ddGTP, ddCTP, ddATP and thermosequenase DNA polymerase in Tris buffer, pH 8.9) and 5 pmol Sp6 primer (5'-TATTTAGGTGACACTATAG-3'). The solution was placed in a thermal cycler and subjected to the following conditions: 94°C for 3 min; then 30 cycles of 96°C for 20 sec, 50°C for 20 sec, 60°C for 4 min. This incorporates fluorescent di-deoxynucleotides into daughter DNA strands.

DNA was precipitated from the sequencing reaction by adding 2 µl sodium acetate (1.5 M), 2 µl EDTA (50 mM, Sigma, UK) and 1 µl glycogen (20 mg/ml, Beckman Coulter, UK) in a final volume of 20 µl, vortexing, and then adding 60 µl 95% (v/v) ethanol (-20°C) and leaving on ice for 10 min. DNA was pelleted by centrifugation at 14,000 g for 15 min at 4°C. The supernatant was removed, the pellet washed with 200 µl 70% ethanol (-20°C), and the tube left open at room temperature until dry. DNA was resuspended in 40 µl de-ionised formamide (JT Baker, USA) for 15 min and transferred to a 96-well plate. The sequences were determined using an automated capillary DNA sequencer (Beckman Coulter, UK) as per the manufacturer's instructions.

2.5 Production of human gene probes

2.5.1 PCR optimisation

Unless otherwise stated, all liquid handling processes in the production of human gene probes were performed by Biomek 2000 Liquid Handler (Beckman Coulter) in 96-well and 384-well format plastic wares (Beckman Coulter, UK; Alpha Laboratories, UK; Millipore, UK).

12 DNA polymerases from 8 suppliers (Table 2.4) were tested for their suitability in PCR amplification of array probes. Each 50 µl PCR reaction consisted of 40 pmol of each HHV-7 U76 primer (Thermal BioSciences GmbH, Germany), 1 µl template DNA (serial dilution from 10 ng to 1 fg), 1.25 units DNA polymerase enzyme, 40 µmol dNTP (Promega), in 5 µl 10× reaction buffer containing 1.5 mM $MgCl_2$. Two sets of reaction mixes were prepared for all enzymes, one of which proceeded to PCR reaction immediately after preparation and the second set was left at room temperature for 3 h before PCR was performed. PCR was performed in a Primus Multiblock thermal cycler (MWG, UK) as follows: 94 °C for *x* min (Table 2.4); 40 cycles of 94 °C for 30 sec, 55°C for 30 sec, 72 °C for 30 sec; then 72 °C for 7 min. Products were checked for size and integrity by agarose gel electrophoresis (section 2.4.4).

Table 2.3 Conditions for DNA polymerases in batch test

			First heating step	
Code	Enzyme Name	Supplier	Temperature °C	Duration (min)
A	Taq	Qiagen	94	3
B	HotStar	Qiagen	94	15
C	Taq	Novagen	n/a	n/a
D	Hot Start	Novagen	95	10
E	Taq	Roche	94	2
F	FastStart	Roche	95	2
G	Taq	Sigma	94	1
H	Taq	Bioline	n/a	n/a
I	Immolase	Bioline	95	7
J	Taq	Promega	94	2
K	DyNAzyme	Finnzymes	94	2
L	YieldAce Hotstart	Stratagene	94	2

2.5.2 Preparing lymphochip clones for human array probes

Human gene probes were generated from a lymphochip gene set of ~2100 human clones (MRC gene service, Appendix B). The clone sets were in bacterial glycerol stocks (LB, 8% glycerol, 50 µg/µl ampicillin) arrayed in 96-well formats. 5 µl of each clone were inoculated into 150 µl 2×LB (50 mg/ml ampicillin) and allowed to grow for 4-8 h at 37 °C with shaking. 5 µl of this initial culture was used to inoculate 800 µl LB (50 mg/ml ampicillin) and grown in the same way for 12-16 h. Bacterial clones were diluted 1:5 with distilled water and heated at 98 °C for 10 min and cooled on ice for 10 min to

break up the cell wall and release plasmid DNA. Cell debris was removed by centrifugation at 1500 g, 3 min to obtain DNA-containing cell-free supernatant for PCR (section 2.6.2).

2.6 Production of custom-made microarrays

2.6.1 Array PCR to amplify viral and human probes

Human array probes for in-house host-pathogen microarrays were produced by PCR amplification using pT7T3 primers (Table 2.3). Each 100 µl PCR reaction consisted of 40 pmol of each primer (Thermal BioSciences GmbH, Germany), 10 µl plasmid DNA-containing supernatant, 40 µmol dNTP (Promega, UK), and 2.5 units Immolase DNA polymerase (Bioline, UK), 10 µl 10× reaction buffer (160 mM $(NH_4)_2SO_4$, 1 M Tris-HVL (pH 8.3), 0.1% Tween-20). Viral probes were amplified in a similar manner to the human probes, using 40 pmol of each pGEM-T primer, 10 ng purified DNA. PCR was performed in a Primus Multiblock thermal cycler (MWG, UK) as follows: 95 °C for 7 min; 40 cycles of 94 °C for 30 sec, 55°C for 30 sec, 72 °C for 60 sec (30 sec for viral probes); then 72 °C for 7 min. Products were checked for size and integrity by agarose gel electrophoresis (section 2.4.4). Single band products were selected and re-distributed onto fresh 96-well plates using Biomek HitPicking Wizard programme (Beckman Coulter, UK).

2.6.2 Human and viral gene probes purification

PCR products for both human and viral genes were transferred into Montage PCR_{384} purification plates (Millipore, UK) for purification. Unincorporated nucleotides were removed by size selective filtration under 10 inches Hg vacuum for 15 min. Purified DNA was resuspended in 20 µl spotting buffer, containing 3× saline sodium citrate buffer (SSC, Sigma, UK) and 1.5 M Betaine (Sigma, UK) and transferred into 384-well arraying plates (Genetix, UK). The process was automated using Biomek 2000 Liquid Handler (Beckman Coulter, UK).

2.6.3 Microarray spotting and processing

The purified DNA probes in spotting buffer were deposited onto aminopropyl silane-coated Nexterion Slide A (Schott, UK) using the Qarray^2 arrayer (Genetix, UK) with a 12-pin print head of Genetix aQu 75um split pins (Genetix, UK). The probes were spotted in triplicate, in a format of 36 sub-grids, with 196 spots (14 columns by 14 rows) per grid. Array spotting was performed in collaboration with Edward Tsao, Royal Free & University College Medical School, London. Spotted arrays were air-dried for 12+ hours, followed by heating at 80 °C in dry oven for 3 h to immobilise the DNA. Unbound surfaces on the arrays were inactivated by immersing in 60 ml blocking solution: 5× SSC, 0.5% (w/v) bovine serum albumin (BSA, Sigma, UK), 0.5% (w/v) non-fat powdered milk (Tesco, UK) at 50 °C for 45 min, followed by washing 2 times in deionised water. Then the arrays were incubated in deionised water at 95 °C for 2 min to denature the DNA. Denatured arrays were washed and blocked again (with fresh solution) at 50 °C for 30 min. Finally the arrays were washed 2 times in deionised water, immersed in isopropanol (Sigma, UK) for 30 seconds and immediately dried by spinning at 200 g, 2 min. Processed arrays were stored in a dry environment for use within 24 hours.

2.7 Measuring transcript abundance using custom-made microarrays

2.7.1 RNA extraction from TRIzol

2.7.1.1 Total RNA extraction

Each ml of frozen TRIzol lysates (section 2.3.7) were thawed at 37°C and centrifuged at 12,000 g, 10 min, 4 °C to remove insoluble material. DNA and proteins were removed from the supernanant by two rounds of chloroform purification, by adding 0.2 ml (first round) and 0.5 ml (second round) of chloroform, and the solution shaken by hand for 15

s. After incubation at room temperature for 2-3 min, the aqueous layer was separated from the organic layer by centrifuging at 12,000 g, 15 min, 4 °C. After the two rounds of chloroform purification, RNA in the aqueous phase was precipitated by incubating with 0.5 ml isopropanol for 10 min at room temperature. Precipitated RNA was pelleted by centrifugation at 12,000 g, 15 min, 4 °C and washed with 1 ml 75% ethanol. Ethanol was removed after centrifugation at 7500 g, 5 min, 4 °C. The RNA pellet was air-dried and resuspended in 50 - 100 µl RNase-free water (BDH, UK). RNA concentration was estimated by UV absorbance at 260 nm, using the ND-1000 Spectrophotometer (Nanodrop Technologies, USA).

2.7.1.2 DNase treatment of total RNA

To ensure only pure RNA was used in the subsequent analysis, total RNA was treated with 5 units of DNase (Promega, UK) per 50 µg RNA at 37 °C for 1 h to remove any contaminating DNA. DNase treated RNA was purified two rounds using an equal volume of 25:24:1 phenol:chloroform:isoamylalcohol (Sigma, UK), followed by one round of chloroform extraction (as in section 2.7.1.1), with vigorous shaking and centrifugation at 16,000 g, 10 min at room temperature to separate the RNA from the reaction components. Purified RNA was then precipitated for 1-2 h at –20 °C by adding $1/5^{th}$ volume ammonium acetate (8M) and 2.5 volume 95 % ethanol (–20 °C), followed by centrifugation at 16,000 g, 30 min, 4 °C. The RNA pellet was washed in 200 µl ethanol (80 %, –20 °C), centrifuged at 16,000 g, 10 min, 4 °C, and air dried for 5 min before resuspension in 100 µl Rnase-free water. Concentration and quality of the DNase treated RNA was accessed using Agilent RNA 6000 Nano assay with the Agilent 2100 bioanalyser (Agilent Technologies, UK). If the RNA was not used immediately it was stored at –80 °C.

2.7.1.3 Messenger RNA purification

Messenger RNA (mRNA) was purified from total RNA using Oligotex (Qiagen, UK), which separates nucleic acids containing poly-dA sequences through hybridisation to oligo-dT bound-microscopic polystyrene-latex particles. Before purification, the Oligotex suspension was warmed to 37 °C, vortexed and then returned to room

temperature and buffer OEB heated to 70 °C. Total RNA was made up to 250 µl and mixed with 250 µl buffer OBB and 15 µl Oligotex suspension. This was incubated at 70°C for 3 min to denature the RNA and then annealed at room temperature for 20 min. Subsequently, the Oligotex was pelleted by centrifugation at 14,000 g for 2 min and the supernatant discarded. The pellet was resuspended in 400 µl buffer OW2, mixed and applied to a Qiagen spin column. This was centrifuged at 14,000 g for 1 min, the Oligotex resuspended in an additional 400 µl buffer OW2 and centrifuged again. The bound mRNA was eluted by resuspending the Oligotex in 200 µl of buffer OEB (70 °C) and centrifuging through the column at 14,000 g for 1 min. This was repeated to give a final volume of 400 µl. The mRNA was concentrated to a volume of 10 – 20 µl by centrifugation through a Microcon Centrifugal Filter Device YM-30 (Millipore Corporation, UK) at 14,000 g for 8 min. The mRNA was recovered by inverting the column and centrifuging at 1000 g for 3 min. Concentration and quality of the mRNA was accessed using Agilent RNA 6000 Nano assay with the Agilent 2100 bioanalyser (Agilent Technologies, UK). Purified mRNA was stored at –80 °C.

2.7.2 RNA extraction from RLT lysate

2.7.2.1 Total RNA extraction

For time course infection samples of a low cell number (1×10^5), total RNA was extracted using the RNeasy mini kit (Qiagen, UK). SupT1 cells were harvested by centrifugation, washed in 2 volumes of PBS, and centrifuged again to remove the PBS. 1×10^5 cells were lysed in 350 µl RLT buffer (containing 0.1% (v/v) β-mercaptoethanol) and homogenised by vortexing. 350 µl 70% ethanol was added to the lysate and the mixture transferred onto spin column. The liquid was pulled through by centrifuging at 8000 g, 15 sec, room temperature. The column was washed once with 350 µl RW1 buffer and centrifuged as above, followed by digestion of DNA using 10 µl DNase, 70 µl RDD buffer (Qiagen, UK) at room temperature for 15 min. The column was washed again with 350 µl RW1 buffer and centrifugation, then was washed twice with 500 µl RPE buffer (diluted 1:5 with 100% ethanol) and centrifuged at 8000 g, room temperature (first time for 15 sec and second time for 2 min). RNA was eluted from the

column in 30 µl RNase-free water by centrifugation at 8000 g, 1 min, room temperature. Total RNA extracted was quantified using the Agilent RNA 6000 Nano assay with the Agilent 2100 bioanalyser (Agilent Technologies, UK) and stored at –80 °C.

2.7.2.2 Sense-strand RNA amplification

Messenger RNA was linearly amplified from 250 ng of DNase treated total RNA using the SenseAmp kit (Genisphere, USA). Firstly, total RNA was annealed to 50 ng oligo-dT24 primer and 500 ng random nanomer primer in 11 µl at 80 °C for 10 min and cooled on ice for 2 min. First strand cDNA was reverse transcribed from the RNA/primer mix using Superscript II reverse transcriptase (Invitrogen, UK), 500 µM of each dNTP, in the presence of 4 µl 5× first strand buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM $MgCl_2$), 10 mM DTT, 1 µl Superase-In RNase inhibitor in a total volume of 20 µl. The reaction was performed at 42 °C for 2 h. First strand cDNA was purified using the MinElute PCR purification kit (Qiagen, UK). The cDNA was bound to the MinElute column in the presence of 500 µl PB buffer. Column was centrifuged at 16,000 g, 1 min; washed in 750 µl PE buffer and centrifuged again. 500 80% ethanol was then passed through the column by centrifuging for 2 min and the column air-dried for 5 min. The purified cDNA was eluted in 10 µl EB by first equilibrating for 2 min at room temperature and then centrifuging for 2 min. In order to allow RNA amplification directly from the purified first strand cDNA (i.e. no second strand synthesis), a poly-dT tail was added to the 3'-end of the first strand cDNA using 2 µl terminal deoxynucleotidyl transferase (Genisphere, USA), 2 mM dTTP nucleotides, 2 µl 10× reaction buffer, in 20 µl total volume. The reaction took place at 37 °C, 3 min, and was quenched at 80 °C, 10 min. A poly-dA-anchored T7 promoter was then annealed to the poly-dT tail by incubating at 37 °C for 10 min. The 3'-end of the cDNA was then 'filled in' to complete the double stranded T7 promoter using 1 µl Klenow enzyme, 400 µM of each dNTP, 1 µl 10× reaction buffer, by incubating at room temperature for 30 min. *In vitro* transcription (IVT) was then performed from the T7 promoter-modified cDNA (12.5 µl) using 2 µl T7 enzyme mix, 8 µl nucleotide mix, 2.5 µl 10× T7 reaction buffer. Amplified RNA was synthesised by IVT reaction at 37 °C for 6 h in a Primus Multiblock thermal cycler (MWG, UK). Amplified RNA was stored at 4 °C until clean

up, using the RNeasy kit (Qiagen, UK), in a similar procedure to section 2.7.2.1, except that 250 µl of 100% ethanol was used instead of 350 µl 70% ethanol; and that the RW1 buffer washing and DNase treatment steps were eliminated. The concentration of the amplified RNA was measured using the Agilent 2100 bioanalyser (Agilent Technologies, UK).

2.7.3 Reference mix

The reference RNA comprised the following cell lines in the stated proportions: U937 (monocytes) 25%, HeLa (epithelial) 25%, SSCEM (T-cells) 15%, Ramos (B-cells) 15%, Huh7 (liver) 5%, MRC5 (fibroblasts) 5%, HHV7-SupT1 (DC) 5%, HHV6A-SupT1 (U1102) 2.5%, HHV6B-SupT1 (Z29) 2.5%. RNA of the individual cell lines were extracted, purified and quantified separately and mixed before RNA amplification or labelling.

2.7.4 Fluorescent labelling of RNA

Fluorescent dye-labelled cDNA was synthesised from 250 – 1000 ng mRNA or amplified RNA (aRNA) using the CyScribe First-Strand cDNA labelling system-dCTP (Amersham Biosciences, UK). Briefly, mRNA or aRNA were annealed to 1 µl each of oligo-dT and random primers, in a total volume of 11 µl, at 70 °C for 5 min, and cooled to room temperature for 10 min. This RNA/primer mix was added to the labelling reagents: 100 units reverse transcriptase, 5 mM DTT, 50 µM Cy5- or Cy3-dCTP, 1µl dNTPs (except dCTP), 4 µl 5× CyScript buffer in 20 µl total volume. Reverse-transcription reaction was performed at 42 °C for 90 min in Primus Multiblock thermal cycler (MWG, UK), to produce fluorescent dye-labelled first strand cDNA.

Labelled cDNA was denatured with 20 µl sodium hydroxide (0.1 M), 5 µl EDTA (0.5 M, pH 8.0) at 70 °C for 10 min, and neutralised with 20 µl hydrochloric acid (0.1M). To suppress hybridisation of repetitive DNA, 3 µg of Cot-1 DNA (Invitrogen, UK) was added. Unincorporated nucleotides and dye were removed by washing with 450 µl

Tris-EDTA buffer (TE) through a Microcon YM-30 filter device (Millipore, UK) by centrifugation at 14,000 g for 6 min. This was repeated 2 more times with 300 µl TE and centrifuging for 8 min. Purified cDNA was collected by inverting the Microcon column and centrifuging at 1000 g for 3 min, to a final volume of 13.5 µl or less.

2.7.5 Hybridisation onto microarray

For each hybridisation, 250 – 1000 ng of Cy5- and Cy3-labelled cDNA samples were mixed with the hybridisation mix to block non-specific binding and denature the DNA: 12 µl 20× SSPE (saline sodium phosphate EDTA buffer, Sigma), 1.1 µl of EDTA (0.5 M, Sigma), 2 µl of poly-dA (40-mer, 8 µg/µl, Thermal Hybaid), 2 µl yeast tRNA (4 µg/µl, Sigma), 1 µl 10% SDS, in TE buffer to a final volume of 45 µl. This mixture was incubated for 2 min at 98 °C to denature the labelled DNA, followed by incubation at 37 °C for 20 min. 1 µl of 100× Denhardt's (Sigma, UK) was added to prevent non-specific binding and the mixture centrifuged at 16,000 g, 15 min to remove any insoluble materials.

Meanwhile the pre-blocked array (section 2.6.3) was placed into a hybridisation chamber and warmed to 65 °C on hot block. A glass coverslip (22 × 64 mm, thickness No. 0, SLS, UK) was washed in isopropanol (Sigma, UK) for 30 seconds and immediately spun dry in empty centrifuge tubes at 200 g for 2 min. When the labelled cDNA was ready 40 µl was pipetted onto centre of the array, over which the coverslip was gently but quickly placed to allow dispersal of liquid by capillary action. 150 µl of 4× SSPE (65 °C) was added to the chamber to maintain humidity and the lid immediately closed. Arrays were hybridised at 65 °C by immersion in waterbath for 16 – 20 h.

The hybridised array was removed from the chamber and immersed in 2× SSPE, 50 °C until the coverslip was detached. The array was then sequentially washed in three wash solutions of increasing stringency at room temperature. 2× SSPE for 2 min, 1× SSPE for 2 min and 0.1× SSPE for 3 min. After the third wash, the array was immediately

spun dry at 200 g, 2 min. The array was either scanned (section 2.8) or stored in the dark.

2.8 Array scanning and data collection

Each array was scanned using a GenePix 4000 microarray scanner and the GenePix Pro 4.0 software (Axon Instruments, USA), at 10 μm resolution. Cy3 and Cy5 were simultaneously excited at 532 nm and 635 nm respectively and the fluorescence measured by photomultiplier tubes (PMTs). The voltages across the PMTs were adjusted (generally set around 500V) so that the signals from the array elements were balanced. The software generated a combined image from the measurements, on which a template was overlaid to identify each spot to their corresponding gene. Array spots from which no signal could be detected were flagged as not found by the GenePix software. The image was also examined by eye to flag any gene spots damaged by debris, scratch or smudging. For each spot the software calculated the Cy5 and Cy3 median intensities of all pixels in the spot, as well as the local background signals. The data were exported to the TM4 software package (TIGR, UK), an open-source program developed for microarray data analysis, for processing and analysis (Saeed et al., 2003).

2.9 Array data processing and normalisation

Array data were converted into *.mev* file by the TIGR ExpressConvert software (TM4, TIGR, UK), and read into the MIDAS software (TM4, TIGR, UK). The program first removed bad quality data already flagged by the GenePix software or manually. It then removed questionable data based on a 'stringent tolerance policy', so that data with a signal-to-noise ratio of <2.0 in any one channel were removed. This was followed by a low intensity filter, in which data with a integrated intensity below 10,000 were removed. Filtered data were then normalised to correct for various systematic and experimental errors. First of all, total intensity normalisation was applied to balance the values in both channels. This was achieved by applying a normalisation factor (f) to the Cy5 signal of each spot using the Cy3 value as the reference, so that:

$$Corrected\ Intensity\ (Cy5) = Original\ Intensity\ (Cy5) / f$$

The normalisation factor was calculated as follows:

$$f = \frac{\sum Intensity(Cy5)}{\sum Intensity(Cy3)}$$

This was followed by the LOWESS normalisation in the global mode, to correct for intensity-dependent local errors. This was performed by applying LOWESS algorithm (Cleveland 1979) with a smoothing parameter of 0.33, to adjust the Cy5 value using the Cy3 value as a reference.

2.10 Clustering analysis

Normalised data from multiple arrays to be compared were assembled and imported into the Mev software (TM4, TIGR, UK) for multiple array comparison. A low percentage filter was set to include only data points that were present in 80% or 100% of the total number of arrays compared. The data were mean-centred by gene (to set the mean ratio of each gene across all arrays to 0), and by array (to set the mean ratio of all gens within one array to 0) to scale the data so that relative comparisons could be made. Genes and arrays were clustered by average-linkage hierarchical clustering, using the uncentred Pearson correlation as the similarity metrics, based on the method developed by Eisen *et al*, 1998.. Clustered results were visualised as a dendrogram in the MeV software.

2.11 Significance Analysis of Microarrays (SAM)

Significance Analysis of Microarrays (SAM), enables binary comparisons of microarray data, a gene-specific t-test that gives significantly regulated genes, and includes a permutation algorithm that allows estimation of error, termed the false discovery rate (FDR) (Tusher et al, 2001). SAM has been incorporated into the TMev software (TM4, TIGR, UK). Genes that passed the 100%-presence filter (i.e. genes that passed SNR filtering across all arrays analysed) were separated into two groups (inhibitor-treated vs

untreated) and their expression signals compared by SAM. The FDR was determined by a user-defined Delta value, which was set at 3% (i.e. 3 genes in every 100 genes found to be significantly regulated may be falsely identified). Lists of significantly up- and down-regulated genes were assembled and used in subsequent hierarchical clustering for visualisation as a dendrogram in TMev.

2.12 Reverse transcription-PCR (RT-PCR)

2.12.1 RT-PCR reaction

Specific HHV-7 transcripts in total RNA, mRNA or amplified RNA were amplified by reverse transcription PCR (RT-PCR) using the OneStep RT-PCR kit (Qiagen, UK). Briefly, 0.05 pg – 50 ng template RNA were reverse transcribed into cDNA by the Omniscript and Sensicript reverse transcriptase enzymes present in the OneStep enzyme mix (Qiagen, UK), in the presence of 0.6 μM HHV-7 ORF specific forward and reverse primers (Table 2.3), 400 μM of each dNTP, and 10 μl 5× reaction buffer, in a final volume of 50 μl. Reverse transcription was allowed to take place at 50 °C for 90 min in Primus Multiblock thermal cycler (MWG, UK). This was followed by a heating step at 95 °C for 15 min, which inactivated the reverse transcriptases and activated the HotStarTaq DNA polymerase which was also present in the OneStep enzyme mix. The newly synthesised cDNA was amplified by PCR with 35 cycles of 94 °C, 30 sec; 55 °C, 1 min; 72 °C, 1 min; followed by a final extension step of 72 °C for 10 min. In control reactions to test for DNA contamination, template RNA was added after the heating step, or the OneStep enzyme mix was replaced by HotStarTaq enzyme (Qiagen, UK). RT-PCR products were analysed by agarose gel electrophoresis as in section 2.4.4.

2.12.2 Gel purification

One volume of the RT-PCR reaction was separated on a 1% w/v agarose gel as in section 2.4.4. The band of the correct size was excised from the gel and DNA purified with the QIAquick gel extraction protocol (Qiagen, UK). 3 volumes of buffer QG (v/w)

were added to the gel slice and incubated at 50°C for 10 min. 1 volume of isopropanol was added, the solution vortexed and then applied to a QIAquick column. The solution was passed through the column by centrifugation (14,000 g, 1 min), washed with 500 µl buffer QG and then with 750 µl buffer PE. DNA was eluted in 30 µl of buffer EB.

2.12.3 Sequence analysis of gel bands

The identity of the DNA extracted from agaorse gels (2.12.2) was confirmed by sequence analysis as detailed in section 2.4.6.3.

CHAPTER 3

PRODUCTION OF HHV-7 MICROARRAYS

3.1 Introduction

3.1.1 Reasons for HHV-7 microarray production

Viruses are intracellular parasites that rely on their continual interaction with the host cellular environment and their exploitation of cellular processes for replication and spread. Understanding these interactions at the molecular level contributes valuable knowledge to virus life cycles, pathogenesis and host defence. Conventional single gene studies provide useful information about the functions of individual genes or proteins, but they do not offer an integrated view on the complex multifactorial infection processes. One approach to study viral-host interaction is by gene expression profiling. Transcription is one of the earliest steps of relaying genomic information to the synthesis of proteins, which are in general the active effectors of many biological processes in both the host and the viral life cycle. Since transcription is a tightly regulated process, changes in transcription during infection usually reflects changes at the protein and functional level (Guerra et al., 2004; Tang et al., 2005). This in turn reveals the biological state and activities in the infected cells (Geiss et al., 2003; Song and Stinski, 2002), or the stages of the viral life cycle (Ahn et al., 2002; Chambers et al., 1999; Jenner et al., 2001; Paulose-Murphy et al., 2001). By applying microarray technology to gene expression profiling, it is possible to monitor the transcriptional changes of not just one gene, but thousands of genes simultaneously (section 1.2). Therefore a global picture of the complicated changes that take place during infection can be obtained.

Microarrays have been used to study global expression changes of cellular genes during infection by viruses, while other studies have monitored changes in viral gene expression (see section 1.2.4). Herpesviruses, in particular, have been the focus of many microarray-based gene expression studies, as the high-throughput technology is ideal for analysing their large DNA genomes. Viral gene expression profiling has been performed on HSV (Stingley et al., 2000), HCMV (Chambers et al., 1999), Kaposi's sarcoma associated herpesvirus (KSHV) (Jenner et al., 2001; Paulose-Murphy et al., 2001), murine herpesvirus type 68 (MHV68) (Ahn et al., 2002) and HHV-6 (Ohyashiki et al., 2005). The microarrays used contain 85 to 180 probes that target the corresponding sequences in the viral genomes. These arrays also contain variable

numbers of human probes, from 4 (Ohyashiki et al., 2005) to 88 (Paulose-Murphy et al.), mainly housekeeping genes for control and normalisation purposes. The probes were produced either by PCR (Ahn et al., 2002; Jenner et al., 2001; Paulose-Murphy et al., 2001) or by oligonucleotide synthesis (Chambers et al., 1999; Ohyashiki et al., 2005; Stingley et al., 2000) before robotic deposition onto membrane (Ahn et al., 2002; Jenner et al., 2001) or glass (Chambers et al., 1999; Ohyashiki et al., 2005; Paulose-Murphy et al., 2001; Stingley et al., 2000) arrays. Apart from one study (Ohyashiki et al., 2005), all viral microarrays were produced in-house. These viral microarray studies have characterised the transcription profiles of herpesviruses over the course of infection *in vitro* and *in vivo*; some have identified particular genes of interest. More detailed summary of the findings is provided in Chapter 5.

The above mentioned studies focussed on either the host or viral gene expression. However, since viral infection is a multifactorial process involving both parasitic and cellular contributions, analysis of both viral and host genes in parallel during infection would be highly desirable. This could be achieved using a DNA microarray with probes to both viral and host genes to measure transcript abundance of viral and cellular genes together. One study has included probes for HIV as well as cellular genes on the custom-made arrays (van 't Wout et al., 2003). However the study only focussed on cellular expression while the HIV gene probes were used as positive controls rather than for transcription analysis.

Human herpesvirus type 7 (HHV-7), with its high prevalence and apparently non-pathogenic nature, is believed to be a virus very well adapted to humans (section 1.1). It is therefore useful to study the transcription profiles of the virus and the host to understand the life cycle and pathogenesis (or the lack of it) of HHV-7 infection. In order to obtain the global pattern of host and viral gene expression, we have made a DNA microarray containing probes that target all predicted HHV-7 genes and a subset of cellular genes. This is the first description of an integrated host-pathogen microarray for HHV-7.

The aim of work described in this chapter was the production and testing of HHV-7 microarrays, forming the foundation for further analysis in the later sections.

3.2 Results

3.2.1 Production of HHV-7 array probes

To determine global expression patterns of human herpesvirus type 7 (HHV-7) genes, a custom-made DNA microarray was created to measure the abundance of HHV-7 gene transcripts. This involved production of the HHV-7 array probes, which were then spotting onto glass-slides by automated robotics. The production of HHV-7 array probes involved a series of DNA amplification, cloning, sequence confirmation and re-amplification, represented on a diagram (Figure 3.1).

3.2.1.1 Primer design and initial PCR amplification from genomic DNA

Array probes were designed to target the 86 predicted open reading frames (ORFs) in HHV-7, positioned towards the 5'-end of each ORF (Figure 3.2). This positional preference was to distinguish between ORFs with co-terminal 3'-sequences, for example U53 and U53.5. As an attempt to detect any previously unidentified genes, 10 probes (X1-X10) were also designed for some of the larger intergenic 'gap regions' in the HHV-7 genome between known ORFs (Figure 3.2). The production of array probes involved initial PCR amplification of DNA segments approximately 300 bp in length using specific primers (Table 3.1) designed using the Primer3 program (Rozen, 2000), based on the following criteria – optimal melting temperature of 52 °C, primer length of 20 bases, and prevention of secondary structures. Genomic DNA extracted from HHV-7 infected SupT1 cells were used as the template for PCR. The amplified DNA segments were analysed by agarose gel electrophoresis (Figure 3.3), which showed that amplification was successful for 85 ORFs and 4 intergenic probes (X1, X3, X5, X7). One ORF (U77) failed to amplify with the initial primers (U77-1) therefore primers were re-designed (U77-2) to successfully amplify the U77 probe (Table 3.1). No further attempts to amplify the six failed intergenic probes (X2, X4, X6, X8, X9, X10) were performed.

Table 3.1 HHV-7 array primers

ORF	Strand[a]	Acc. No.	Forward primer (5'-3')	start[b]	Reverse primer (5'-3')	start[b]	probe size[c]
DR1	+ (2)	AAC40718	CTGCTGCAACCACAGAACAT	5	GCGAGCTTGTGCTTGTCTCT	306	302
DR6	+ (2)	AAC40719	AGAAATGCTCCGCGCTGT	9	TGAAGTTCCAGGCTCCAGTT	320	312
U2	-	AAC40720	CCACAAAAGGACAGCAGACA	759	CAGATTCTCCGCTTTTGGAT	1076	318
U3	-	AAC40721	CAAGTTCAAACCCCTCTGGA	863	GCAGATTTGGGGTCTGACTT	1172	310
U4	-	AAC40722	TGAAGAAACGCGAGAAAACC	1319	TGTTTGTAAAGGGCCGCTAA	1609	291
U7	- (2)	AAC40723	GGTCCGGTTCTCGGTAATTT	2409	TGTTCCTGTGTGCAATGATG	2695	287
U8	-	AAC40724	AATTGGCACGTCTTTCCTTC	769	TTGAGACGATGGCAGAACAA	1079	311
U10	+	AAC40725	AGCGGAGCGTTCGAAATAA	16	CTGGTGACGCGACAATAAAA	328	313
U11	-	AAC40726	TGCCCAGTTTAAGCAAATCTT	1852	GGATTTCGGATGAAGCAAAG	2220	369
U12	+ (2)	AAC40727	TACGCCAACAGTACGGAAAG	159	CTGTATGTGCAGGACGCAC	442	284
U13	+	AAC40728	CAACAGCCATCTAATTGCAGA	13	AAAATGGTGACGAACGAACTT	292	280
U14	+	AAC40729	CGAAAACTGCAACTTTGTGC	45	TTCATCGAATCGTCGCATAG	360	316
U15	- (3)	AAC40730	TGTGTGAATTGAGCGACGAT	421	GAAACCTGGAGAAGACAACGA	719	299
U17	- (2)	AAC40731	GCAGCTGCTGTTGTTACGTC	695	TAGGCATGCACTCGCAAAT	1036	342
U18	-	AAC40732	TGTTAGCCCGAGAAGCAAA	531	CAAATTTGAAATGTCACCAGAA	830	300
U19	-	AAC40733	AAGCTGACACCCATTCTTCG	708	ATGTGGCTTAACGACGATGC	967	260
U20	-	AAC40734	TCAGCGTTAAAATTGTGGTGAT	848	TGTGGCACATAAGTTGACAATG	1123	276
U21	-	AAC40735	TCGCTTCATCAGTCCTTTTGT	961	ATCTTGCTGTTTTGCGTTCC	1284	324
U23	-	AAC40736	TGCTGCAATTAAAATCACAAGG	127	GTACGCCCACTCTGGAAATG	452	326
U24	-	AAC40737	TTCGTACCCCCATATCGAAC	8	TGCATCAGGAAAATCTATCACC	176	169
U24A	-	AAC40738	AGGGATATTTAGAGCACCTGA	3	CAACCATTTTTTCCCAGAGAC	155	153
U25	-	AAC40739	CAAATGCCCTCTGTCACCTT	679	AGCGAATTTGTGCTGTGTCA	952	274
U26	-	AAC40740	CCATTGAATGCGTTGGAAAT	549	GGGCCTTGTTTTAGGGAATG	853	305
U27	-	AAC40741	CTGATGCCGTGACTAAAAACC	696	TCGTAGCAACAGAGAACACCA	1090	395
U28	-	AAC40742	ATGGAATCAAATTCCGAAGC	2120	AAAACATGGGTCTCTTGAGGA	2390	271
U29	-	AAC40743	CTCCAAACAAAGACACCTTTTC	477	TTTGCTAGCTTTTGGTGTTCC	772	296
U30	+	AAC40744	TGCGGCTGAAAGTATCACAG	108	AATAGGCGATGCCAATTCCG	427	320
U31	+	AAC40745	GCAGGAAGTACAAATCAAAACG	13	TTGTTTCGCCAAAATTAGAACT	280	268
U32	-	AAC40746	CCGGTAAAGAGATGATACTTG	847	TGACATCACACTGCGATACG	1136	290
U33	-	AAC40747	CCTTGCATTGCCTGTACAAA	119	ATGGTTCGTACAGTCATTTGC	371	253
U34	-	AAC40748	CGCATCTCGGTTTCGTCTA	361	TCCCGCAGACTTTCGTGT	715	355
U35	-	AAC40749	ATCTCCGAGTGACTGCATTG	95	AACGACAATGCCAAAACAGA	309	215
U36	+	AAC40750	TGGAACTCCGATTCCTTTTC	16	GGGAGTTCTAGGTGCAGCAT	327	312
U37	+	AAC40751	ACACGTAGACTGAGACGAGCA	16	TGCCACCATAGCATCTGAAG	339	324
U38	-	AAC40752	ACCAGATGGAATAACGAAATGT	2704	ATGGATCTGGTGTCCTTCTTT	3042	339
U39	-	AAC40753	TTGGCATTGGCAAAGTTGTA	2124	AGCTGACTTTGTCATGACTGGA	2407	284
U40	-	AAC40754	GCTTGCATTCATTGAGTTGG	1827	TTGCAATCATTGTGTGTCTTG	2157	331
U41	-	AAC40755	CAGGCTTTCGATTTCCGTTA	3012	CAATATGTACCGCAGCATGG	3362	351
U42	-	AAC40756	CCGGAATACTTGCGTAAACC	1217	GTATTAGGGCGGCAACGATA	1524	308
U43	-	AAC40757	CACTTCCTTCTGGAACATGAGA	2256	TGTGTTTGCAACGGAATATGA	2575	320
U44	+	AAC40758	CTACCAGATGCTCTATTCTGAGC	51	TTTTCATGAACACAACATCCA	343	293
U45	-	AAC40759	CCGTTTTGAGGAATTCCAAG	863	TGCAAAGTCTACCCTCCAGAA	1139	277
U46	+	AAC40760	GCCAATTATATTGCTGGCATTT	27	CAGTTTCTTTGGCGAGAGTCTT	247	221
U47	-	AAC40761	TAATGTCCATCGCGGAAAAT	616	CAATGCTAGTGTTCACGTTGAT	923	308
U48	-	AAC40762	TTGTATGCGAAACAGGTTTGA	1722	AAATGGATGGAAACACTGGAA	2032	311
U49	+	AAC40763	GCCCCCTGTTAGAAGACGTA	18	TCTGTTGCTTGTCGTAGTTGC	368	351
U50	+	AAC40764	ATGTACAGACAGAATGCCAGA	206	CCAGCTTCTAAATCGGACAC	619	414
U51	+	AAC40765	CAAATTGGAAATTACTGGCAGA	20	TGATACGTTCCAGTGCCAAC	316	297
U52	-	AAC40766	GCTGAGCTGCATTTTCTTCA	435	TCGAACAACTCAACCTCCAA	738	304
U53	+	AAC40767	GTGGCAGGCTTCCTATGTGT	16	GGATCTGGTGGCAGAGTTTT	308	293
U53.5	+	AAC40768	AGCTAGTCTTCCCGCTCCTC	27	TCATCATCCGAATCTGTTGG	323	297
U54	-	AAC40769	TGCGCTCAACTATTTGGACA	1026	TGGAATCCTGTTATGTGGTCTG	1364	339
U55A	-	AAC40770	ACCCTTTTCAAGATAGGCAAA	910	TGCAGCCAGAACTCAATTTT	1220	311
U55B	-	AAC40771	GGAAAAACTCCCATCTGCTC	941	CAAAGAAGAATTCCAGGGTAAA	1240	300
U56	-	AAC40772	GGCAGAATCAGCTCCCAGT	500	AACTCACTAATGCCGTCATCC	818	319
U57	-	AAC40773	GGCGCAGTTACGTCGTTT	3713	CACAGCTGAAATTTTCCCAAA	4024	312
U58	+	AAC40774	TCACTGTGCTAAAAATGTTCCA	54	TATATAGAACCGGGCCTTCG	355	302
U59	+	AAC40775	ACAATGAGACAGCAACTGTGC	53	AAAAACAGTTTGGGCTTTTCA	323	271

ORF	Strand[a]	Acc. No.	Forward primer (5'-3')	start[b]	Reverse primer (5'-3')	start[b]	probe size[c]
U66	-(2)	AAC40780	ACCGCAGATCCATCTTTGAA	4704	TGCAATTCAGAAAGCGTATCA	5085	382
U62	+	AAC40776	AAATTTGTCTGAACTGTGATTTC	77	ATTTCCACCGCGAGCGTAT	176	100
U63	+	AAC40777	ATGTCAGTCAATCACAGTTGC	30	ACGGTCTCCATATTTGGCTCT	299	270
U64	+	AAC40778	GCTTGTATATACGGAATGGAC	156	GTATGCGAGGTGCTCCAGA	406	251
U65	+	AAC40779	CCACGAATATCGTTGTGTGG	93	GGCGTAATAGGTCGCACATT	389	297
U67	+	AAC40781	GCTGCAATTTATAAAGAAACGAC	22	AGCGATGTTCTCACTGGTCA	343	322
U68	+	AAC40782	GTCACTGCATGAATTAATAAAAC	3	GTCGGTGACGGATTTCACTT	303	301
U69	+	AAC40783	GCAGCTTAAGACACCCCAAA	6	CATGCACTCAATCCATCTCG	311	306
U70	+	AAC40784	TGGCAATAGATTACGCACAAA	2	CGCATCTTGTAAGCGGACAT	303	302
U71	+	AAC40785	ATGGGCTCAAAGTGCTGTAA	1	TTCTTCTGGGTTAACGAAACG	200	200
U72	-	AAC40786	TCCCCCATCCAAGTTATATCC	674	TGGCATTGAGCAGAGTAGATG	1040	367
U73	+	AAC40787	TTTTTGGAATGGTTTGGACA	31	GACGCTCTTTAATTTCGCTGT	328	298
U74	+	AAC40788	ATGCTGAGACTCACTCCTTG	104	GGCATGTACAAAACCGCATT	392	289
U75	-	AAC40789	CTCAGCCGAATCTCTTCTAGC	482	ACTAAACAAACACGCCGAACA	694	213
U76	-	AAC40790	TGCTAAGGGTGTCAGCGTAA	1365	AGGGACAATCTATACACAGTG	1703	339
U77-1	+	AAC40791	GAATTTGCTCAACCGTTGCC	217	TCTGCGATGACGTTCCAATAT	541	325
U77-2	+	AAC40791	GCAAAACAAGCAGCATTCAA	283	CACTTTGGCTTTCCCATTTG	587	305
U79	+ (3)	AAC40792	TGATTGCGGAGGATAGAGAAT	2	CCTTTCCCTCGTTTTTCTGA	322	321
U81	-	AAC40793	TGCTGAAAGCTAATCCGTGA	453	GGATGCTGGAAAACATTTCA	749	297
U82	-	AAC40794	CCAGTTGTGCCCTGATTCT	397	ATTTCTTATCTCCATTCTAAACC	716	320
U84	-	AAC40795	TTGTTGCAAAGTTTGTCGTTG	633	AACTGTTTGTAAGGCCAAAGG	913	281
U85	-	AAC40796	CCAGCATCGTTAATTGTTCG	512	TGCCTTTCTTAGGTGCAATTT	842	331
U86	-	AAC40797	TTTCTCCACTTGCTTGAGGAT	3318	GCACTACAGCTGACGTGGTT	3614	297
U90	- (3)	AAC40798	CGAGGAAGCTCCTTCTTCAA	3405	TGGACAACCATCTGAAACCAT	3723	319
U91	+ (2)	AAC40799	AAACGTGTATCAAGGCCAAAA	22	CATCCGGGTTGTTTACTGCT	345	324
U95	+	AAC40800	ATGGAGAGTGGGGATAGTTTC	1	ATGTTCCTTGGTTTCCTTGC	289	289
U100	- (10)	AAC40801	TGGAATGCAGAGTGTTGTTTC	2322	CCAATACGGTCCCACTTGA	2645	324
DR1'	+ (2)	AAC40802	CTGCTGCAACCACAGAACAT	5	GCGAGCTTGTGCTTGTCTC	306	302
DR6'	+ (2)	AAC40803	CTGTTCAGCTCCAGCCAAG	23	TGAAGTTCCAGGCTCCAGTT	320	298
X1	n/a	n/a	CGAGATCGTAACGAACTTG	1	CAATGGTCACTGAGCCATCA	271	272
X2	n/a	n/a	TTAGCGTCATCATACAACGG	1	GCTTACTTTGCGGAAGTGGG	203	204
X3	n/a	n/a	GTCTGCGTTACTTCTGTCAC	1	TATCTCCGTGACAAAGCATAC	301	302
X4	n/a	n/a	CAACCTGATAGAAGTAGGGG	1	TTTGGGAGTCAAGCAAGCAC	264	265
X5	n/a	n/a	CAGTTGAATGAGGACTGCAG	1	CAACCTTCAGTAGTGTGTGG	229	230
X6	n/a	n/a	AGGCAACCTAAAACTGGCTG	1	CAGCGAAAGAGGTGAAGATTG	339	340
X7	n/a	n/a	ATTTCATGGGCCCTGAAACC	1	ATGCAGTAGCATACTACAATCC	344	345
X8	n/a	n/a	TGTGTGCAGTTTGAGAACATG	1	TTTCTTATATTGCTAGCCCTCC	400	401
X9	n/a	n/a	AGGAACTGCGGTTGCCTTC	1	GGGTGAGAACCCATAAAACC	359	360
X10	n/a	n/a	CAGTTAGTCTGCTGGAAAGC	1	TCCAAACCTACGCAATGTATC	300	301
pGEM-T			GCATGGCGGCCGCGGGAATT		GGCGGCCGCGAATTCACTAG		
T7T3			TTGTAAAACGACGGCCAGTG		GGGAATTTGGCCCTCGAGG		
SP6 promoter primer			TTCTATAGTGTCACCTAAAT				
T7 Promoter primer			TAATACGACTCACTATAGGG				

[a] Strand +, ORF in same direction as genome; -, ORF on anti-strand; number in brackets indicate number of predicted exons if >1

[b] Position of the first base of primer, relative to first base of the corresponding HHV-7 ORF (in the direction of genome); in the case of ORFs on the complementary strand, the forward primer is closer to the 3'-end of the ORF; for integenic probes, the start position is always 1

[c] Size of array probe in bp

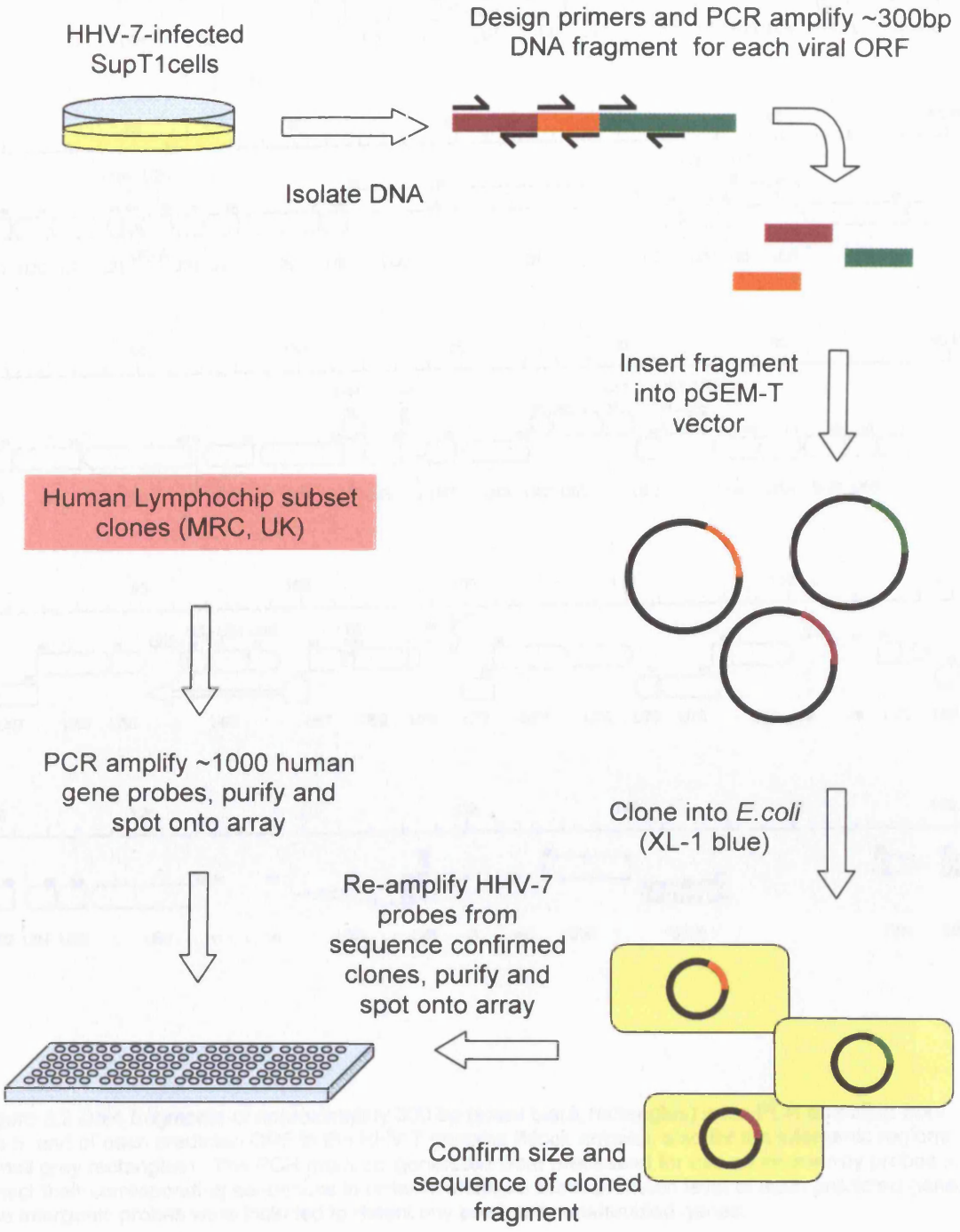


Figure 3.1 HHV-7 microarray production

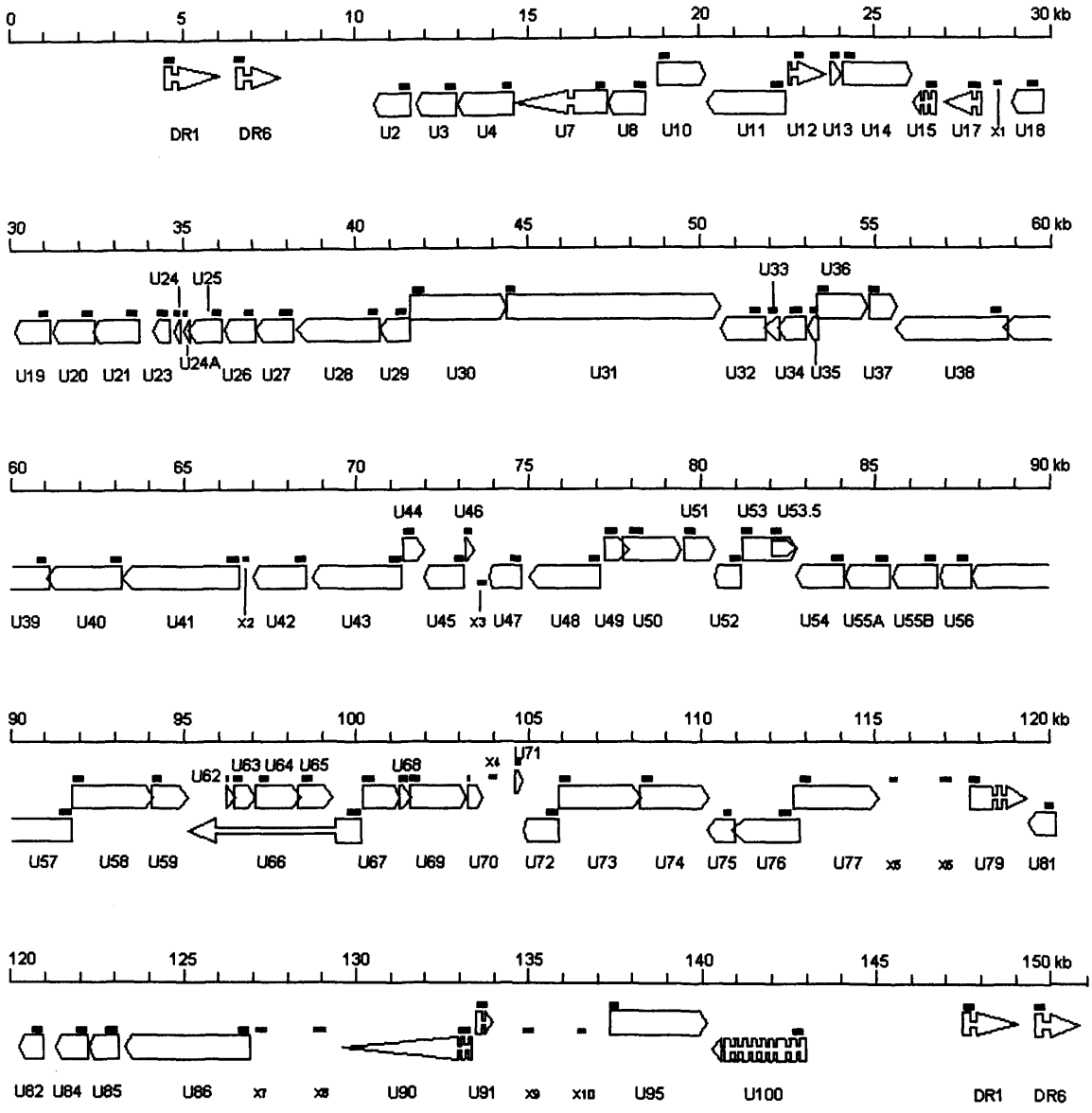


Figure 3.2 DNA fragments of approximately 300 bp (small black rectangles) were PCR amplified from the 5'-end of each predicted ORF in the HHV-7 genome (block arrows); also for ten intergenic regions (small grey rectangles). The PCR products generated were processed for use as microarray probes to detect their corresponding sequences in order to measure the expression level of each predicted gene. The intergenic probes were included to detect any previously unidentified genes.

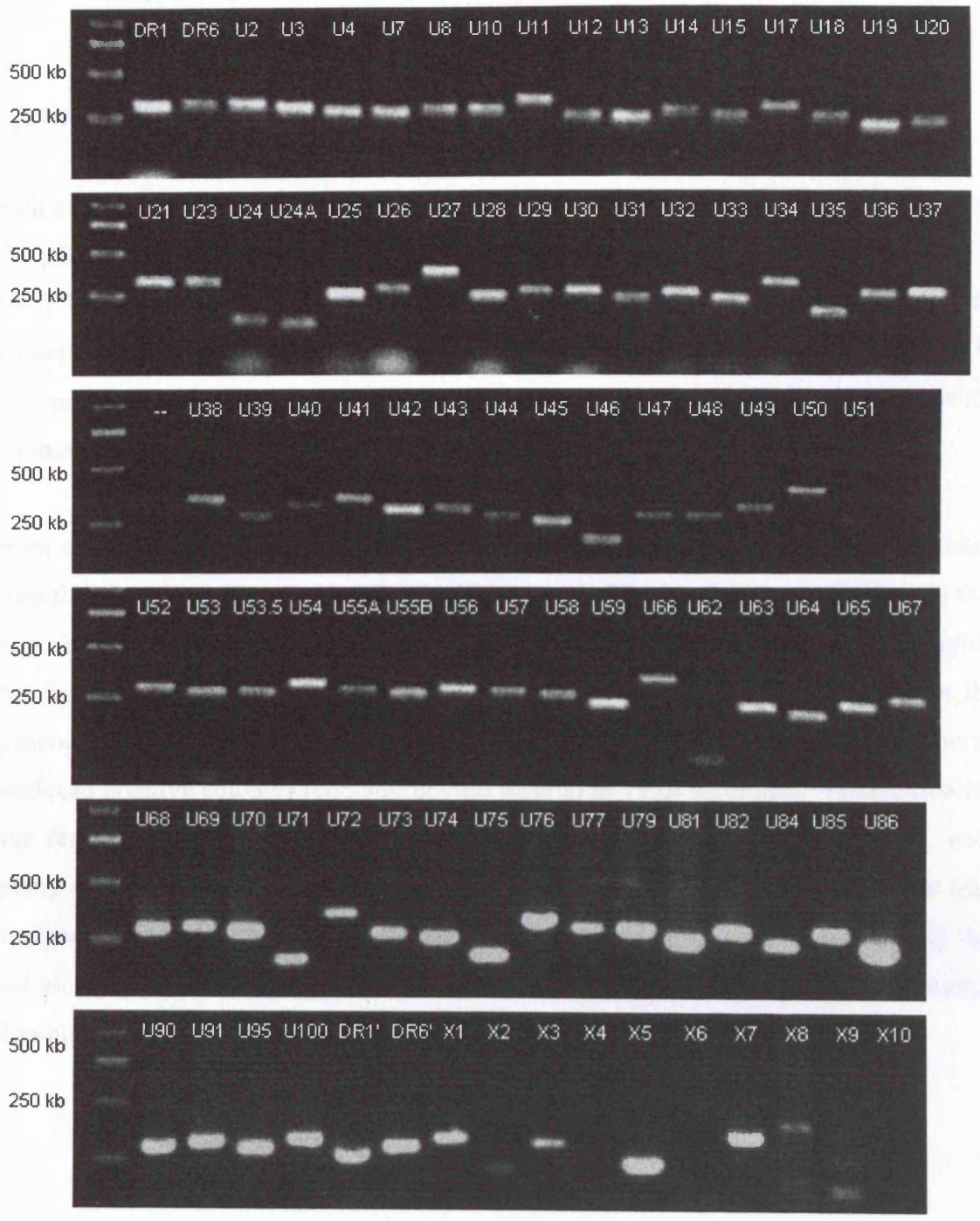


Figure 3.3 PCR amplification of HHV-7 fragments using specific primers (Table 3.1) and DNA extracted from HHV-7–infected SupT1 cells.

3.2.1.2 Cloning of array probes

Each amplified DNA segment was cloned into the bacterial vector pGEM-T by ligation. The pGEM-T plasmid was transformed into competent bacteria *Escherichia coli* (XL1-Blue) by heat shock. Bacteria were plated onto selective agar plates containing ampicillin, X-gal (5-Bromo-4-chloro-3-indolyl b-D-galactoside) and IPTG (Isopropyl b-D-1-thiogalactopyranoside) to identify clones with the viral DNA inserts (white colonies).

From each transformation, 5 random (or all if less than 5) white colonies were picked from the plate for PCR screening for the DNA insert. Universal primers (Table 3.1) that target the pGEM-T plasmid sequences flanking the insertion site were used to amplify the viral DNA insert. PCR products were analysed on agarose gels to check for the presence and approximate size of the inserts (Figure 3.4). The first round of cloning produced positive colonies (containing viral inserts) in 71 of 90 clones. Transformation was repeated for the remaining 19 clones, which produced satisfactory results, each giving at least 2 positive colonies. Finally, out of the 90 clones screened, all had at least two positive colonies apart from 8 clones (A12, B10, B11, C3, C5, D3, E11, F2) that had only one positive colony. 64 clones had the insert in all colonies tested. Therefore cloning was successful for all 90 probes (86 ORFs and 4 intergenic) (Table 3.2).

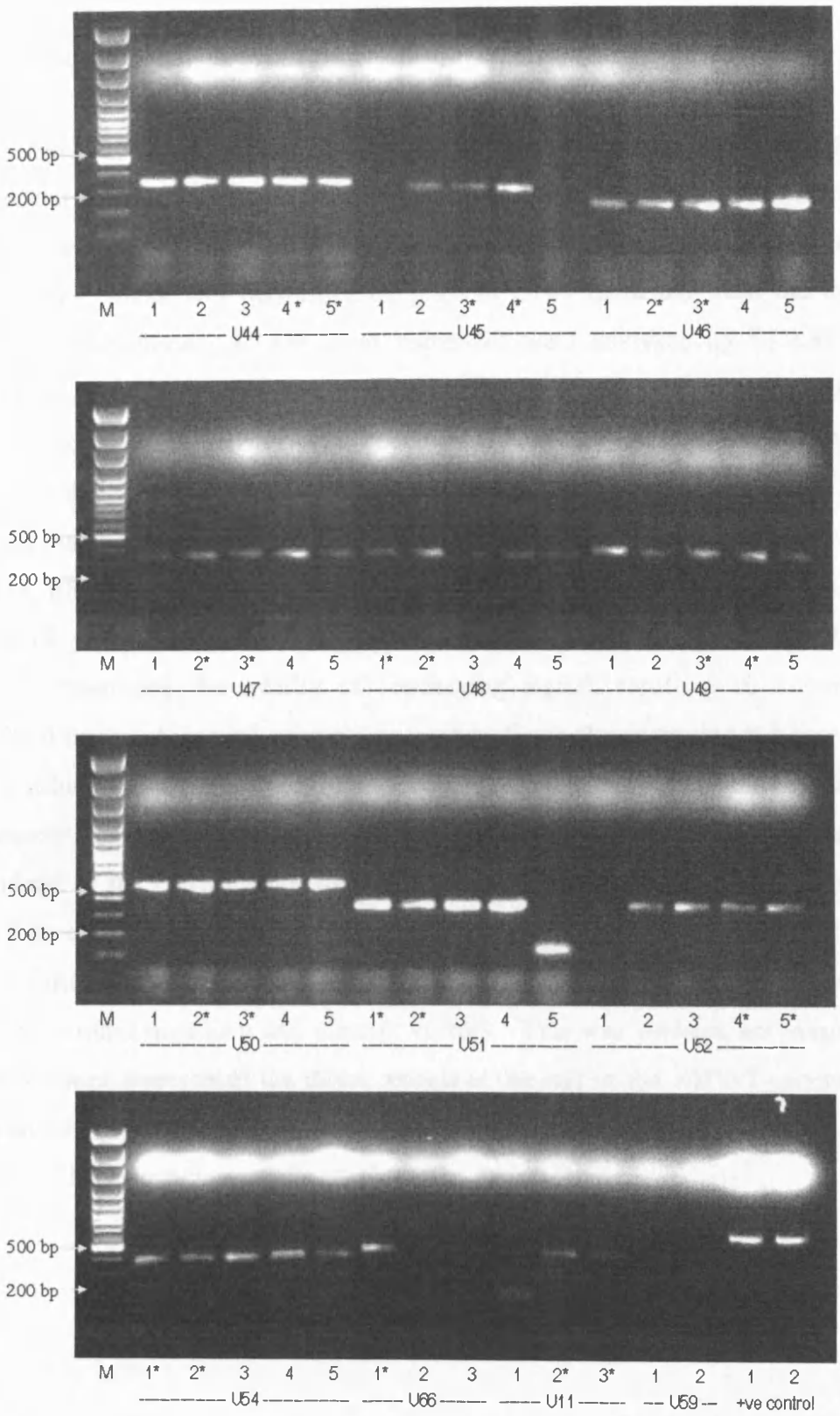


Figure 3.4 Examples of HHV-7 array clones. PCR amplified HHV-7 fragments (Figure 3.3) were cloned into E.coli via pGEM-T vector. Presence of the viral inserts were confirmed by PCR screening using universal pGEM-T primers. Clones marked with * were selected for further amplification into array probes (Figure 3.5) for printing onto microarrays.

3.2.1.3 Sequence verification of clones

PCR screening of clones (section 3.2.1.2) confirmed the presence of DNA insert in all 90 clones but the use of universal primers was not able to confirm the sequences of the inserts. To ensure the clones contained the correct DNA probe, sequence verification was necessary, which was performed on plasmid DNA extracted from the expanded cultures for each clone. All 90 insert sequences were analysed by BLASTN 2.2.2 (Altschul et al., 2001) and were shown to match their corresponding ORFs (or intergenic region) in the HHV-7 RK genome (Megaw et al., 1998), as well as the JI genome (Nicholas, 1996), with adequate percentage identity scores (Table 3.2). All sequences matched with significant p-values, ranging from 1^{-15} to 1^{-128}. A few clones (C12, D4, E8 and G3) initially matched their respective ORFs (U39, U43, U57 and U79) with lower p-values of 1^{-2} to 1^{-7}, probably due to the short sequences (92 – 134 bases) generated and the quality of sequencing signal, resulting in a number of unidentified bases. A second attempt to sequence these clones produced better signals, higher p-values and identities (Table 3.2), confirming that the poorer blast match in the first sequencing attempt was due to sequencing errors, and the insert DNA in the clones were indeed of the correct sequences. Most sequences produced significant matches only to their corresponding region in HHV-7 sequences, with the exceptions A2 and H1, (HHV-7 DR6' and HHV-7 DR1, respectively) which generated many matches to sequences in other organism and cloning vectors. This was perhaps not surprising as these two clones represented the direct repeats at the end of the HHV-7 genome. This was taken into account in the microarray analyses of the four probes DR1, DR6, DR1' and DR6'. The sequencing results are shown in Appendix A.

Table 3.2 Cloning and sequence verification of HHV-7 array probes

Clone	ORF	No. tested[a]	No. with insert[a]	% +ve[a]	Blast match	Genome (RK) location[b]	% Identity[b]	p-value[b]
A1	HHV7DR1	5.0	4.0	80.0	HHV7DR1	4568-4475	93.8	1.0E-29
					HHV7DR1'	147614-147721	93.8	1.0E-29
A2*	HHV7DR6	2.0	2.0	100.0	HHV7DR6	6517-6813	94.8	8.1E-53
					HHV7DR6'	149563-149859	94.8	8.1E-53
A3	HHV7U2	3.0	3.0	100.0	HHV7U2	11501-11634	93.6	1.0E-18
A4	HHV7U3	2.0	2.0	100.0	HHV7U3	12639-12950	95.9	3.4E-57
A5	HHV7U4	5.0	3.0	60.0	HHV7U4	14328-14578	88.7	7.8E-36
A6	HHV7U7	3.0	3.0	100.0	HHV7U7	17036-17323	99.3	2.0E-56
A7	HHV7U8	3.0	3.0	100.0	HHV7U8	18198-18397	78.6	5.4E-17
A8	HHV7U10	3.0	2.0	66.7	HHV7U10	18889-19151	84.4	1.9E-33
A9	HHV7U11	3.0	2.0	66.7	HHV7U11	22076-22443	97.8	1.6E-72
A10	HHV7U12	3.0	3.0	100.0	HHV7U12	22723-23008	100.0	7.7E-57
A11	HHV7U13	3.0	3.0	100.0	HHV7U13	23754-24033	99.6	4.5E-55
A12	HHV7U14	3.0	1.0	33.3	HHV7U14	24152-24456	86.8	5.3E-45
B1	HHV7U15	2.0	2.0	100.0	HHV7U15	26484-26737	95.8	8.0E-45
B2	HHV7U17	5.0	5.0	100.0	HHV7U17	27687-28028	100.0	1.7E-69
B3	HHV7U18	5.0	5.0	100.0	HHV7U18	29464-29764	99.7	8.1E-60
B4	HHV7U19	3.0	3.0	100.0	HHV7U19	30872-31043	100.0	8.8E-30
B5	HHV7U20	2.0	2.0	100.0	HHV7U20	32102-32378	97.9	1.7E-52
B6	HHV7U21	5.0	5.0	100.0	HHV7U21	33382-33706	99.7	3.4E-65
B7	HHV7U23	3.0	3.0	100.0	HHV7U23	34249-34574	100.0	7.2E-66
B8	HHV7U24	5.0	4.0	80.0	HHV7U24	34751-34920	100.0	1.2E-30
B9	HHV7U24A	3.0	3.0	100.0	HHV7U24A	34997-35129	93.4	2.4E-18
B10	HHV7U25	1.0	1.0	100.0	HHV7U25	35880-36107	88.9	7.5E-33
B11	HHV7U26	1.0	1.0	100.0	HHV7U26	36757-37062	98.4	2.1E-59
B12	HHV7U27	6.0	3.0	50.0	HHV7U27	37915-38152	92.8	1.4E-38
C1	HHV7U28	5.0	5.0	100.0	HHV7U28	40404-40677	99.6	3.0E-53
C2	HHV7U29	5.0	5.0	100.0	HHV7U29	41184-41479	100.0	4.3E-59
C3	HHV7U30	2.0	1.0	50.0	HHV7U30	41689-41990	98.7	2.3E-58
C4	HHV7U31	5.0	5.0	100.0	HHV7U31	44412-44679	99.3	5.3E-52
C5	HHV7U32	1.0	1.0	100.0	HHV7U32	51536-51711	98.9	2.9E-31
C6	HHV7U33	3.0	3.0	100.0	HHV7U33	51950-52194	95.1	9.3E-41
C7	HHV7U34	3.0	3.0	100.0	HHV7U34	52573-52927	99.7	5.7E-72
C8	HHV7U35	3.0	2.0	66.7	HHV7U35	53115-53335	95.1	9.9E-38
C9	HHV7U36	5.0	5.0	100.0	HHV7U36	53354-53665	99.0	1.3E-61
C10	HHV7U37	2.0	2.0	100.0	HHV7U37	54812-55136	100.0	1.2E-65
C11	HHV7U38	4.0	4.0	100.0	HHV7U38	58286-58625	99.7	1.0E-68
C12**	HHV7U39	4.0	4.0	100.0	HHV7U39	60748-61031	100.0	5.3E-56
D1	HHV7U40	4.0	4.0	100.0	HHV7U40	62895-63205	95.5	5.0E-56
D2	HHV7U41	4.0	3.0	75.0	HHV7U41	66235-66585	100.0	1.6E-71
D3	HHV7U42	1.0	1.0	100.0	HHV7U42	68232-68520	100.0	1.6E-57
D4	HHV7U43	2.0	2.0	100.0	HHV7U43	71003-71231	99.6	4.4E-43
D5	HHV7U44	5.0	5.0	100.0	HHV7U44	71417-71709	100.0	2.0E-58
D6	HHV7U45	5.0	3.0	60.0	HHV7U45	72845-73120	98.9	3.2E-53
D7	HHV7U46	5.0	4.0	80.0	HHV7U46	73180-73400	99.5	1.1E-41
D8	HHV7U47	5.0	5.0	100.0	HHV7U47	74477-74785	93.2	8.0E-52
D9	HHV7U48	5.0	5.0	100.0	HHV7U48	76761-77073	97.2	1.9E-59
D10	HHV7U49	5.0	5.0	100.0	HHV7U49	77252-77598	97.7	4.8E-67
D11	HHV7U50	5.0	5.0	100.0	HHV7U50	77966-78398	99.8	1.2E-89
D12	HHV7U51	5.0	3.0	60.0	HHV7U51	79591-79828	72.4	7.3E-17
E1	HHV7U52	5.0	4.0	80.0	HHV7U52	80842-81148	92.9	1.1E-51
E2	HHV7U53	5.0	4.0	80.0	HHV7U53	81195-81487	99.7	5.2E-58
E3	HHV7U53.5	5.0	5.0	100.0	HHV7U53.5	82110-82344	95.5	1.9E-40
E4	HHV7U54	5.0	5.0	100.0	HHV7U54	83762-84099	88.3	7.2E-52
E5	HHV7U55A	5.0	5.0	100.0	HHV7U55A	85057-85368	98.7	6.7E-61

Clone	ORF	No. tested[a]	No. with insert[a]	% +ve[a]	Blast match	Genome (RK) location[b]	% identity[b]	p-value[b]
E6	HHV7U55B	5.0	5.0	100.0	HHV7U55B	86455-86754	100.0	5.3E-60
E7	HHV7U56	5.0	3.0	60.0	HHV7U56	87358-87676	98.1	1.3E-61
E8	HHV7U57	5.0	5.0	100.0	HHV7U57	91585-91696	98.2	3.0E-16
E9	HHV7U58	5.0	5.0	100.0	HHV7U58	91873-92109	75.9	2.3E-17
E10	HHV7U59	5.0	5.0	100.0	HHV7U59	94114-94391	95.0	1.1E-49
E11	HHV7U66	3.0	1.0	33.3	HHV7U66	99899-100167	74.4	3.1E-23
E12	HHV7U62	5.0	5.0	100.0	HHV7U62	96323-96424	99.0	7.1E-15
F1	HHV7U63	2.0	2.0	100.0	HHV7U63	96475-96748	91.3	8.9E-45
F2	HHV7U64	2.0	1.0	50.0	HHV7U64	97214-97465	99.6	1.0E-48
F3	HHV7U65	5.0	5.0	100.0	HHV7U65	98461-98727	86.8	2.2E-36
F4	HHV7U67	2.0	2.0	100.0	HHV7U67	100238-100507	80.8	3.8E-29
F5	HHV7U68	2.0	2.0	100.0	HHV7U68	101255-101557	100.0	1.1E-60
F6	HHV7U69	5.0	5.0	100.0	HHV7U69	101606-101911	97.7	1.8E-58
F7	HHV7U70	2.0	2.0	100.0	HHV7U70	103244-103546	97.4	1.6E-57
F8	HHV7U71	5.0	4.0	80.0	HHV7U71	104621-104822	100.0	7.3E-38
F9	HHV7U72	2.0	2.0	100.0	HHV7U72	105539-105905	100.0	3.9E-75
F10	HHV7U73	2.0	2.0	100.0	HHV7U73	105953-106251	100.0	9.0E-60
F11	HHV7U74	2.0	2.0	100.0	HHV7U74	108338-108584	93.1	7.6E-41
F12	HHV7U75	2.0	2.0	100.0	HHV7U75	110843-110684	97.5	1.9E-26
G1	HHV7U76	5.0	5.0	100.0	HHV7U76	112268-112599	96.4	1.7E-61
G2	HHV7U77	5.0	5.0	100.0	HHV7U77	112950-113253	99.0	7.3E-60
G3	HHV7U79	5.0	5.0	100.0	HHV7U79	117734-118055	100.0	1.7E-64
G4	HHV7U81	5.0	3.0	60.0	HHV7U81	119867-120164	100.0	1.5E-59
G5	HHV7U82	5.0	5.0	100.0	HHV7U82	120746-120968	90.2	2.9E-33
G6	HHV7U84	5.0	5.0	100.0	HHV7U84	121970-122251	96.2	1.7E-51
G7	HHV7U85	5.0	3.0	60.0	HHV7U85	122810-123140	98.8	3.1E-65
G8	HHV7U86	5.0	4.0	80.0	HHV7U86	126633-126930	100.0	4.5E-59
G9	HHV7U90	5.0	5.0	100.0	HHV7U90	133050-133371	97.5	1.0E-61
G10	HHV7U91	5.0	2.0	40.0	HHV7U91	133560-133833	99.1	2.5E-64
G11	HHV7U95	7.0	7.0	100.0	HHV7U95	137406-137672	70.4	1.1E-17
G12	HHV7U100	5.0	2.0	40.0	HHV7U100	142646-142948	79.6	8.7E-36
H1*	HHV7DR1'	5.0	3.0	60.0	HHV7DR1	4503-4790	96.2	1.0E-128
					HHV7DR1'	147549-147836	96.2	1.0E-128
H2	HHV7DR6'	5.0	3.0	60.0	HHV7DR6	6351-6828	99.7	3.8E-59
					HHV7DR6'	149577-149874	99.7	3.8E-59
H3	X1 (U17 < U18)	5.0	5.0	100.0	U17< U18	28411-28683	100.0	6.7E-54
H4	X2 (U41 < U42)	n/a	n/a	n/a	n/a	n/a	n/a	n/a
H5	X3 (U46 < U47)	5.0	5.0	100.0	U46 < U47	73522-73822	100.0	1.9E-60
H6	X4 (U70 < U71)	n/a	n/a	n/a	n/a	n/a	n/a	n/a
H7	X5 (U77 < U79)	5.0	4.0	80.0	U77 < U79	115444-115682	100.0	1.0E-43
H8	X6 (U77 < U79)	n/a	n/a	n/a	n/a	n/a	n/a	n/a
H9	X7 (U86 > U90)	5.0	5.0	100.0	U86 > U90	127115-127459	99.4	2.1E-69
H10	X8 (U86 > U90)	n/a	n/a	n/a	n/a	n/a	n/a	n/a
H11	X9 (U91 < U95)	n/a	n/a	n/a	n/a	n/a	n/a	n/a
H12	X10 (U91 < U95)	n/a	n/a	n/a	n/a	n/a	n/a	n/a

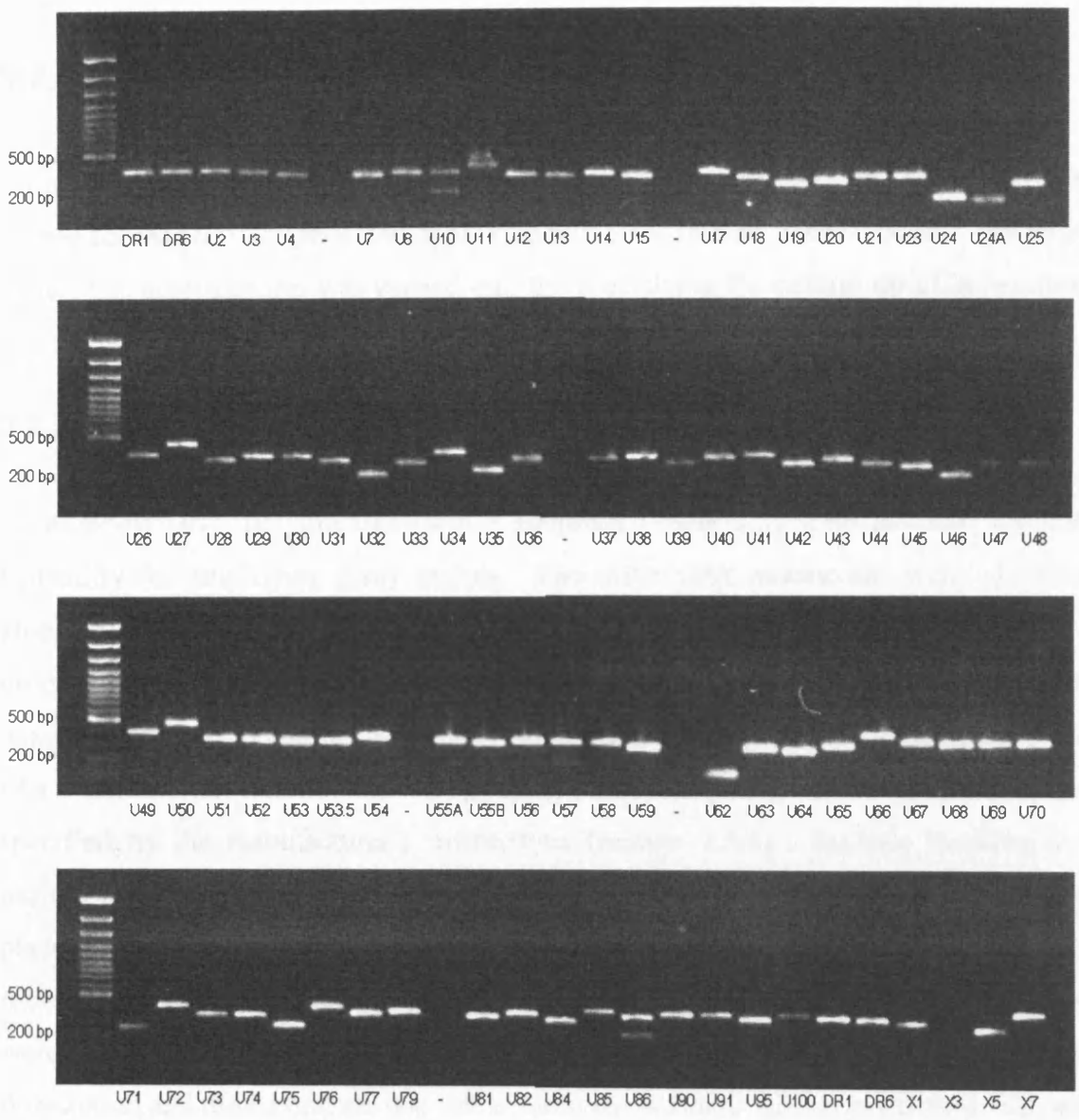
a No. tested, number of colonies used in PCR screening for viral insert; No. with insert, number of colonies containing viral insert; % positive, percentage of colonies containing viral insert out of total number tested

b BLAST results: genome location, location of array probe sequence in the HHV-7 (RK) genome (Megaw et al, 1998); % identity and p-value; calculated by BLAST

3.2.1.4 Amplification and purification of HHV-7 array probes

Using optimised procedures (section 3.2.2), the 90 HHV-7 probes for array spotting were PCR amplified from the sequence verified clones using universal pGEM-T primers (Table 3.1). Therefore the 96-well PCR plate contained 90 probes and 6 blank wells. The products were purified by filtration and resuspended in spotting buffer. All probes were successfully amplified and purified, as shown by agarose gel electrophoresis (Figure 3.5a). It can be seen that despite extensive PCR, purification and robotic liquid handling, the negative controls showed there was no apparent cross contamination. Probes U10 and U86 displayed a doublet, which could be due to cross contamination by another clone, as a common primer was used for all clones. PCR was therefore repeated for these two probes (Figure 3.5b)

a)



b)

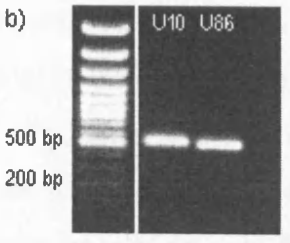


Figure 3.5 a) HHV-7 array probes were re-amplified by PCR from sequence verified clones and purified by filtration. - , negative controls (reaction without DNA template) were included. b) Re-amplification was repeated for probes U10 and U86.

3.2.2 Optimisation of array probe amplification

To produce the in-house host-pathogen microarrays, it was necessary to amplify array probes for 90 HHV-7 genes and approximately 1000 human genes. Before this large-scale PCR amplification was carried out, the procedures for scaling up PCR reactions and the use of extensive robotic liquid handling were first optimised.

3.2.2.1 Optimisation of PCR reaction

12 different DNA polymerases from 8 suppliers (Table 3.3) were assessed for their suitability for amplifying array probes. The criteria for assessment were 1) robust amplification across a range of target concentration; 2) absence of additional PCR product bands; 3) ability to maintain activity after 3 hours at room temperature. Each enzyme was used to amplify a fragment of approximately 300 bp from template DNA of a dilution series from 10 ng to 1 fg, using the same cycling conditions except where specified by the manufacturer's instructions (section 2.5.1). Robotic handling was desired for the preparation of a large number of PCR reactions, but this could only take place at room temperature, hence the enzymes were tested for their ability to withstand room temperature for a prolonged period of time. To do this, two sets of reaction mixes were prepared for all enzymes. One set proceeded to PCR reaction immediately after preparation and the second set was left at room temperature for 3 hours before PCR was performed.

Results (Figure 3.6) showed that when used with our primers, templates, and experimental conditions, enzymes A, E and L produced inconsistent yield, enzymes C and K produced low yield and enzyme G gave no product at all. Enzymes B and C had the lowest sensitivity whereas enzymes F and J had the highest sensitivity and yield amongst those tested. However, these two high-yield enzymes also gave rise to smaller and fainter bands in addition to the product; and enzyme J produced inconsistent product size. In general enzymes H and I gave good yield and clean products. The 3-hour-wait at room temperature reduced the sensitivity of enzyme H, whereas enzyme I was not affected. Therefore enzyme I (Bioline Immolase) was chosen for the consistency, high yield, sensitivity and clean product. The fidelity of the enzymes was

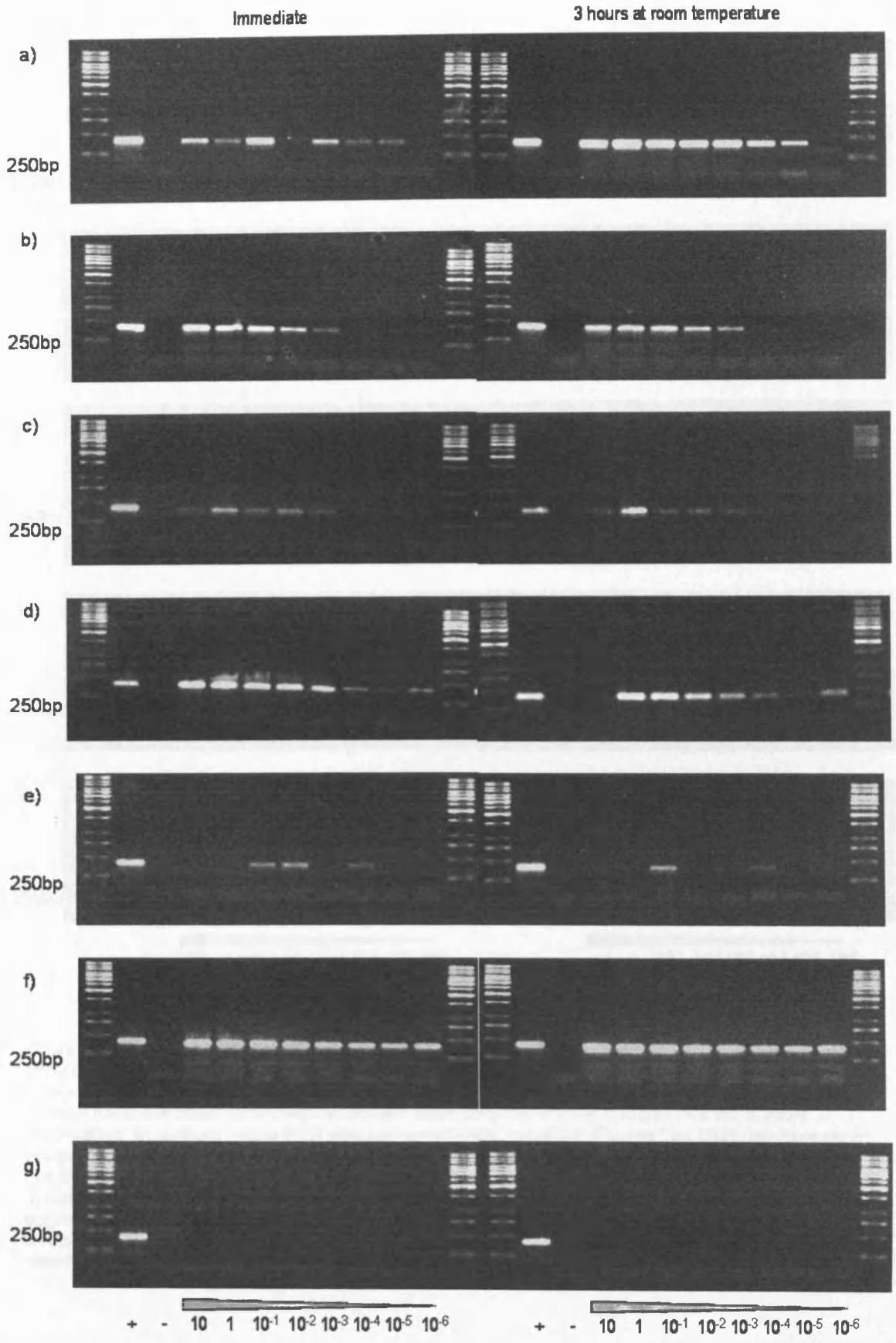
not tested, as the sequences of the human clones were unknown. Any human genes found to show interesting expression patterns will be sequence verified subsequently.

Table 3.3 DNA polymerase batch test

Enzyme Label	Supplier	Name of enzyme
A	Qiagen	Taq DNA polymerase
B	Qiagen	HotStarTaq DNA polymerase
C	Novagen	NovaTaq DNA polymerase
D	Novagen	NovaTaq Hot Start DNA polymerase
E	Roche	Taq DNA polymerase
F	Roche	FastStart Taq DNA polymerase
G	Sigma	Taq DNA polymerase
H	Bioline	BIOTAQ DNA polymerase
I	Bioline	IMMOLASE DNA polymerase
J	Promega	Taq DNA polymerase
K	Finnzymes	DyNAzyme EXT DNA polymerase
L	Stratagene	YieldAce Hotstart DNA polymerase

3.2.2.2 Robotic liquid handling

A Biomeck 2000 liquid handler (Beckman Coulter, UK) was used to quickly and reproducibly prepare PCR reaction plates from master mixes of reagents; to aspirate small volume of PCR products for agarose gel electrophoresis; to select good PCR products, re-organise them onto clean plates and to purify these probes (section 3.2.3.1) for microarray printing.


Immediate
3 hours at room temperature
a)
250bp
b)
250bp
c)
250bp
d)
250bp
e)
250bp
f)
250bp
g)
250bp
+ - 10 1 10-1 10-2 10-3 10-4 10-5 10-6
+ - 10 1 10-1 10-2 10-3 10-4 10-5 10-6

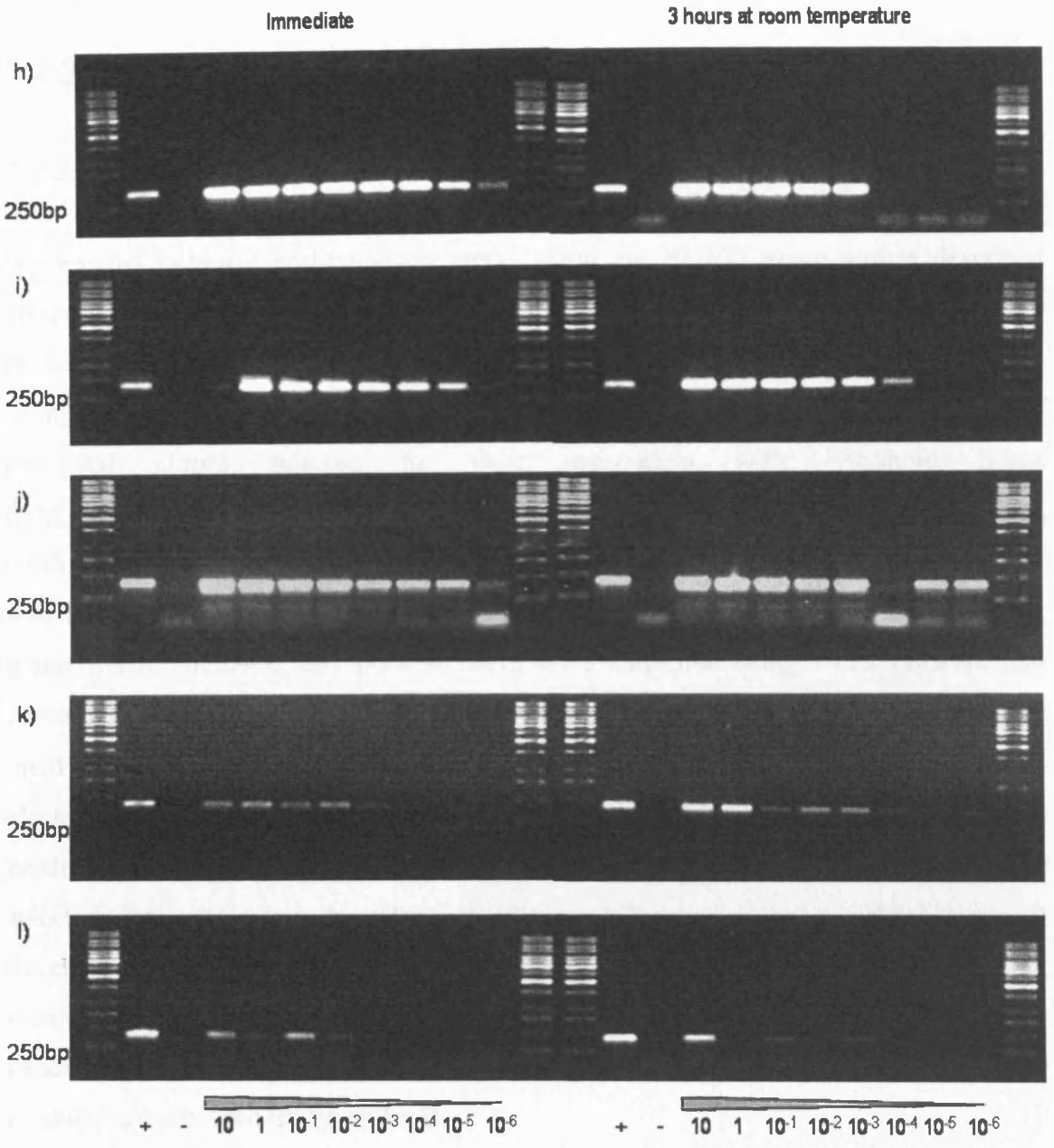


Figure 3.6 DNA polymerase samples from different suppliers were tested for their suitability for use in array probe amplification PCR. Two sets of reaction mixes were prepared for each enzyme, according to manufacturer's guidance, using the same primers and diluted DNA template for all enzymes. One set was run immediately after preparation (left panel) and the second reaction mix left at room temperature for 3 hours before PCR was performed (right panel). a) Qiagen Taq DNA polymerase; b) Qiagen HotStarTaq DNA polymerase; c) Novagen NovaTaq DNA polymerase; d) Novagen NovaTaq Hot Start DNA polymerase; e) Roche Taq DNA polymerase; f) Roche FastStart Taq DNA polymerase; g) Sigma Taq DNA polymerase; h) Bioline BIOTAQ DNA polymerase; i) Bioline IMMOLASE DNA polymerase; j) Promega Taq DNA polymerase; k) Finnzymes DyNAzyme EXT DNA polymerase; l) Stratagene YieldAce Hotstart DNA polymerase. Template DNA dilution in ng; +, positive control; -, negative control (no DNA)

3.2.3 Production of HHV-7–human microarrays

3.2.3.1 Production of human gene probes

We created in-house host-pathogen arrays using the HHV-7 array probes described above (section 3.2.1) as well as human gene array probes. The human gene array probes were produced by PCR amplification from The I.M.A.G.E. Consortium Human Lymphochip subset [IRAY, IRBA, IRAZ, IRBB] (MRC gene service, UK), a set of rearrayed clones selected for their association with lymphoid tissue (http://www.geneservice.co.uk/products/image/image_subsets.jsp). This clone subset consisted of several libraries cloned in various vectors, of which the vector pT7T3-PacI was used for the majority of clones. 1452 clones from the plates IRAY7-IRAY17, plates IRBA7-IRBA10, and plate IRBA12 were amplified using T7T3 Forward and Reverse primers (Table 3.1). PCR was successful for 907 clones (63%), giving good yield (approximately 10-20 μg per 100 μl reaction) and clean single bands on gel electrophoresis analysis of the PCR products. The remainder either failed to produce a product or resulted in multiple bands (Figure 3.7a). The 907 good products were selected for purification by filtration (Figure 3.7b), using a hit-picking protocol developed on the robotic system. Together this means that the PCR amplification, selection of good products and purification were carried out by optimised robotic liquid handling procedures (section 3.2.2). The identities of all 907 human gene array products are provided in Appendix B.

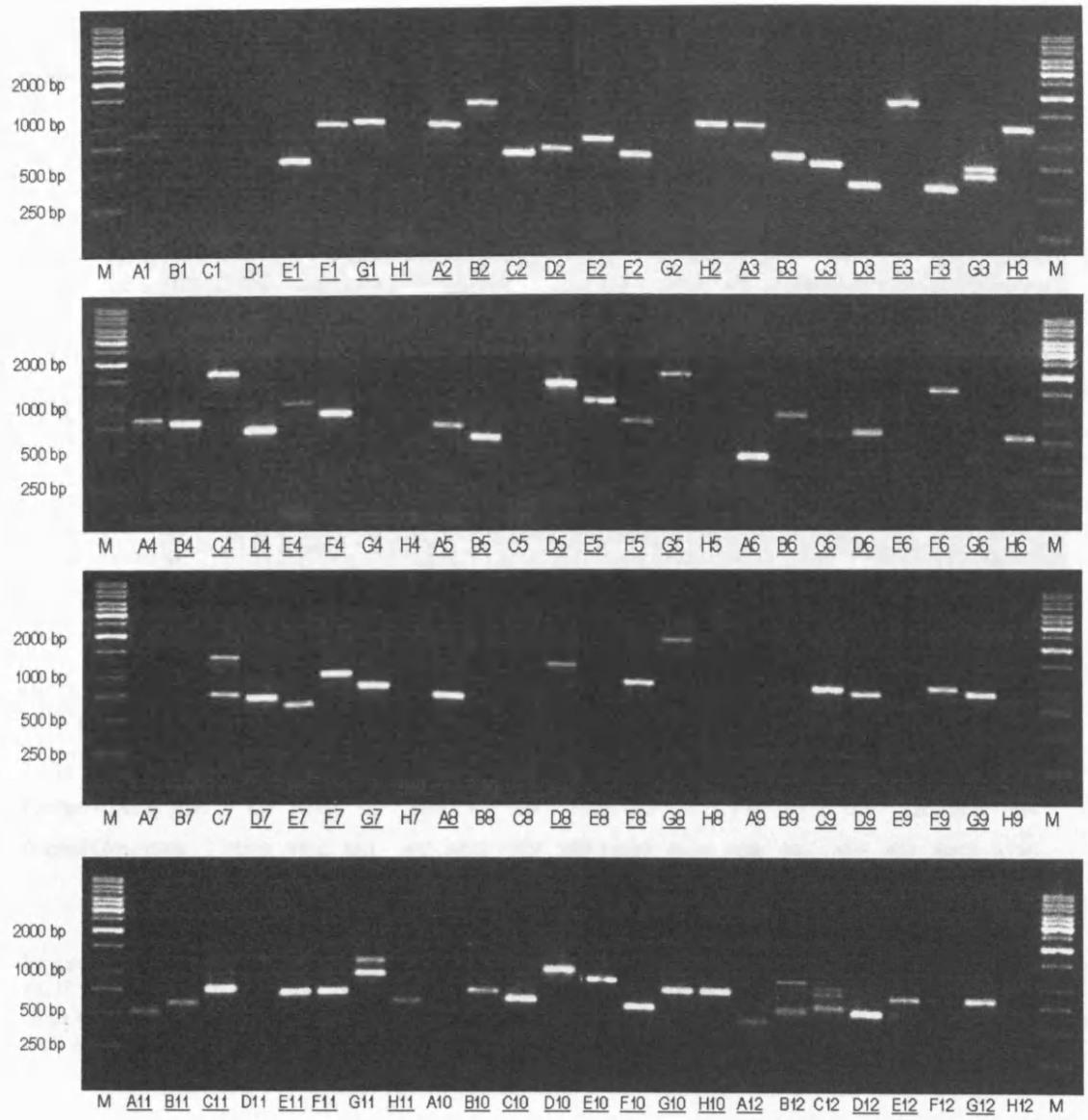


Figure 3.7a An example (plate IRBA7) of PCR amplified products from the Human Lymphochip subset. 68 out of 90 clones (75%, underlined) produced good, clean, single bands and were selected for purification and use as array probes.

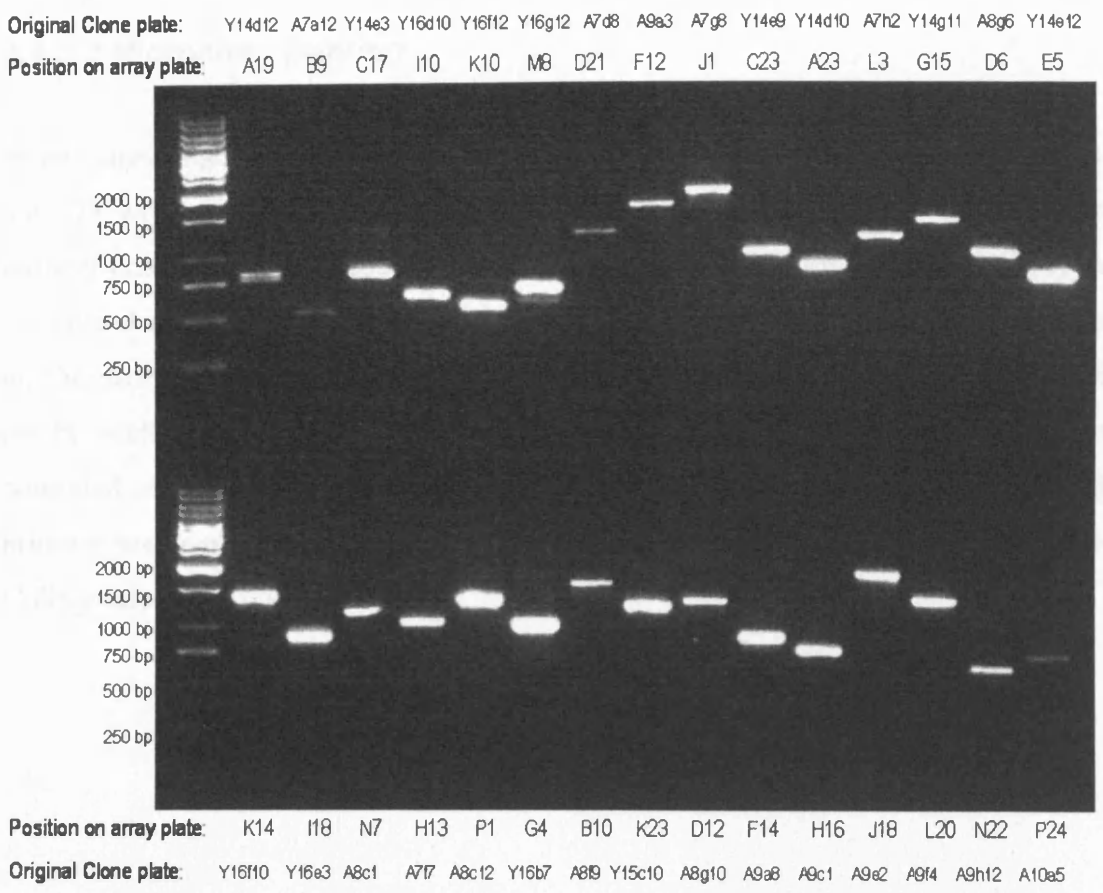


Figure 3.7b Random sample of purified PCR products from the Human Lymphochip subset. Good PCR products were selected, purified by filtration, and analysed by agarose gel electrophoresis. For data tracking purposes, each purified probe were catalogued by their positions on the arraying plate as well as on the original clone plate.

3.2.3.2 Microarray printing

96 custom-designed HHV-7 probes (including 6 blanks) and 907 human probes (section 3.2.3.1) were printed in triplicate (287 spots of HHV-7 and 2721 spots of human probes) onto aminosilane-coated microarray slides by a 12-pin print head (3×4) using automated robotic deposition. Probes for HHV-6 and KSHV genes were also included on the array for the use of other projects. The arrays were laid out in 36 sub-grids (4×9), each of which was a square of 14 columns by 14 rows. In total the array consisted of 7056 spots, some of which were empty (Figure 3.8). The microarray printing was performed in collaboration with Edward Tsao, Royal Free & University College Medical School, London.

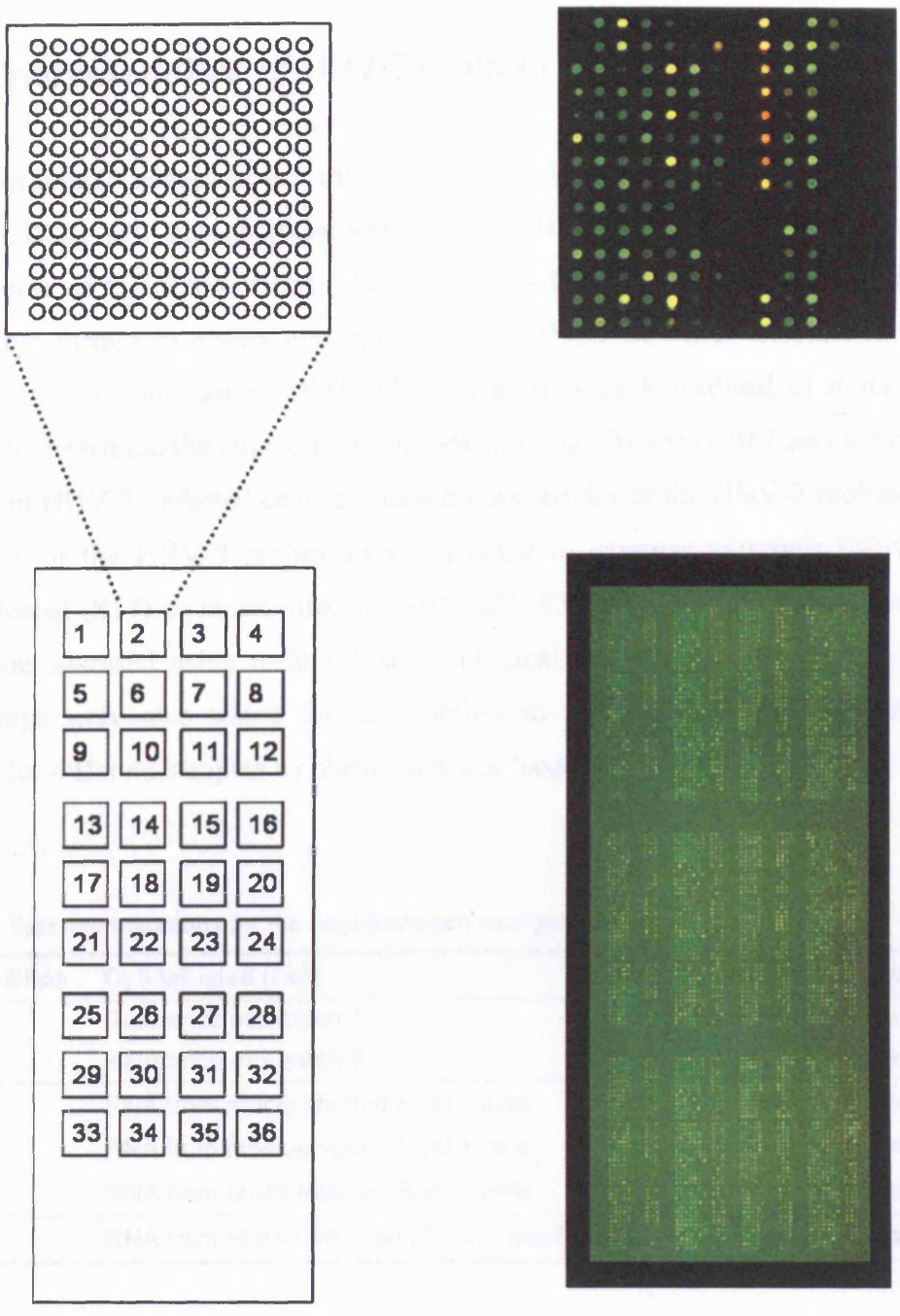


Figure 3.8 Format of HHV-7 microarray. Each array consisted of 36 sub-grids, with 196 spots (14 columns and 14 rows) per sub-grid. Each probe is spotted in triplicate, so that a probe in sub-grid 1, column 1, row 1 was also present in sub-grids 13 and 25, in the corresponding column and row.

3.2.4 Characteristics of HHV-7–human microarrays

The custom-made host-pathogen microarrays were hybridised in various combinations (Table 3.4) to test the characteristics of the HHV-7 probes, as well as for the performance of the whole array. 1) Two arrays (S14 and S20) were hybridised in reverse-dye format to assess dye bias and the effects of different batches of RNA preparation. 2) Three arrays (S02, S29 and S30) were hybridised to mock-infected samples to determine the cut-off point of data filtering. 3) Array S17 was hybridised to RNA from HHV-7–infected cells to check the sensitivity of the HHV-7 probes. 4) The specificity of the HHV-7 probes were evaluated by comparing arrays hybridised to virus-infected (S17) or mock-infected (S02, S29, S30) samples. 5) Reproducibility of arrays was assessed using technical and biological repeats (section 3.2.4.5). 6) The microarrays were also tested for their ability to produce distinguishing expression patterns for different samples by cluster analysis (section 3.2.4.6).

Table 3.4 Test hybridisations for the host-pathogen microarrays

Array Slide	Cy5-labelled (red)	Cy3-labelled (green)
S14	Reference mix batch 1	Reference mix batch 2
S20	Reference mix batch 2	Reference mix batch 1
S02	RNA from mock-infected SupT1 cells	Reference mix batch 2
S29	RNA from mock-infected SupT1 cells	Reference mix batch 2
S30	RNA from mock-infected SupT1 cells	Reference mix batch 2
S17	RNA from HHV-7 infected (72 hpi) SupT1 cells	Reference mix batch 1

The microarray experiment uses a comparative hybridisation approach. In this study we used the two-colour mode, measuring fluorescent intensities of samples labelled with two fluorescent dyes (red Cy5-dye and green Cy3-dye). PolyA-tailed mRNA samples to be analysed on microarrays were extracted from cells, amplified and reverse transcribed with red (Cy5) or green (Cy3) fluorescent-labelled dCTP into labelled-cDNA. The two species of labelled cDNA (Cy5 and Cy3) were co-hybridised onto each array. After washing to removal of unbound cDNA, arrays were scanned to measure the fluorescence intensities at 635 nm (Cy5) and 532 nm (Cy3) to generate an array image,

which was then analysed by image analysis software GenePix Pro 4.0. A matrix (the *.gal* file) defining the position and identity of all the array spots was overlaid onto the image, allowing the calculation of the signal and background intensities (both Cy5 and Cy3 channels) of each spot. These numerical data were used to generate Cy5/Cy3 ratios. Throughout this thesis arrays were hybridised to a Cy3-labelled common reference, so that the Cy3 signals acted as a constant control and the Cy5/Cy3 ratios were thus comparable between arrays. The common reference was a mixture of RNA extracted from 9 populations of cells, composed in the following proportion: U937 (25%), HeLa (25%), CEM-SS (15%), Ramos (15%), Huh7 (5%), MRC5 (5%), HHV7-SupT1 (5%), HHV6A(U1102)-SupT1, (2.5%) and HHV6B(Z29)-SupT1 (2.5%) cells. The purpose of using a mixture of RNA from different cell lines was to include many different RNA transcripts, so that more array probes will be bound and produce fluorescent signals, which in turn enables accurate Cy5/Cy3 ratios to be deduced for as many spots as possible. The inclusion of HHV-7 and HHV-6 RNA in the reference was to produce a signal for viral probes in the Cy3 channel, thus allowing the calculation of Cy5/Cy3 ratios.

3.2.4.1 Effects of Dye-bias and RNA preparation

The reference was made by first mixing the total RNAs in their corresponding proportions, followed by RNA amplification to give enough amplified RNA (aRNA) for many arrays. During the course of the thesis the first batch of reference exhausted hence a second batch was made using the same total RNA sources. It was therefore necessary to ensure there was no significance difference between the two batches of references. To assess the differences between the two batches of RNA reference mix, Ref-1 and Ref-2, they were each labelled in Cy5 and Cy3. Array S14 were hybridised to Cy5-labelled Ref-1 and Cy3-labelled Ref-2; the labelling was reversed for array S20 (Table 3.4).

If the two reference batches gave identical signals on the array, we would expect the median Cy5/Cy3 ratios of all spots on each array to be equal to 1 (or $\log_2$ ratios = 0) and the Pearson correlation values between the two channels should also be 1. Results show that good correlations were indeed obtained. For S14 (Cy5/Cy3 = Ref-1/Ref-2), median

$\log_2$(Cy5/Cy3) = -0.417; Pearson = 0.979 and for S20 (Cy5/Cy3 = Ref-2/Ref-1), median $\log_2$(Cy5/Cy3) = -0.068; Pearson = 0.984 (Figure 3.9). Hence the median $\log_2$(Cy5/Cy3) value was close to zero in S20 and not so close in S14; whereas the Pearson correlation values were both close to and less than 1. To show that the high Pearson correlation values were indeed indicative of the high degree of similarity between Ref-1 and Ref-2 RNA batches, data from array S17 (different species of RNA in the two channels, see Table 3.4) gave a lower correlation of 0.82, indicative of more variation between the RNAs, as expected.

From this observation some conclusions can be made: 1) the Cy5- and Cy3-labelled species (Ref-1 and Ref-2) were highly correlated, as shown by the high Pearson correlation, hence there was negligible variation between the two batches of RNA references; 2) the deviation of the median $\log_2$(Cy5/Cy3) values from zero in S14 may be due to a bias towards Cy3, as a result of variations in labelling efficiency, hybridisation efficiency, or fluorescence detection, a phenomenon well documented for Cy3; 3) S20 has a median $\log_2$(Cy5/Cy3) value closer to zero hence is less affected by dye-bias than S14; 4) normalisation should be applied to remove dye-bias effect when arrays were used for HHV-7 expression profiling. In addition, the distribution of $\log_2$ Cy5/Cy3 ratios around zero on the M/A plots (Figure 3.9) for S14 and S20 was an important observation for subsequent analyses. This was because many subsequent normalisation methods were based on this assumption.

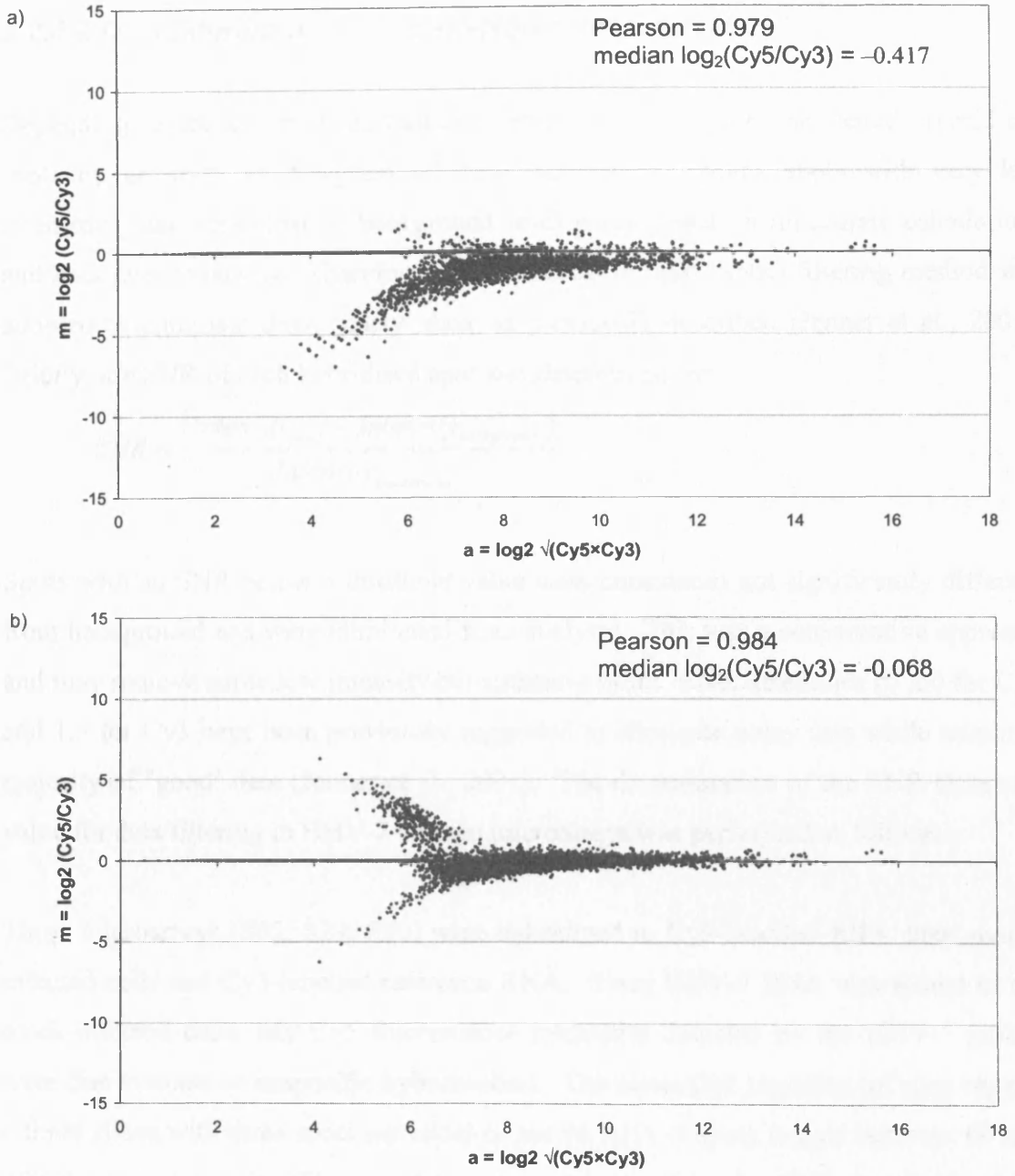


Figure 3.9 Two batches of reference RNA mixture (Ref-1 and Ref-2) were each labelled in Cy5 and Cy3 and hybridised to each other in reverse-dye format. Figure shows m/a plots of two arrays, a) Slide S14 hybridised with Cy5-labelled Ref-1and Cy3-labelled Ref-2; b) Slide S20 hybridised with Cy5-labelled Ref-2 and Cy3-labelled Ref-1. The $\log_2$ (Cy5/Cy3) ratios of each spot is plotted against the geometric mean intensity of that spot. Pearson correlation values between the two labelled-species in each array are shown.

3.2.4.2 Data filtering and Signal-to-Noise Ratio threshold

Depending on factors such as transcript abundance and probe sequence, hybridised spots on an array gave signals of varied intensities. Some spots with very low intensities that were close to background level would result in inaccurate calculations and data interpretations. Therefore, a signal-to-noise ratio (SNR) filtering method was adopted to eliminate these 'noisy' data, as previously described (Jenner et al., 2001). Briefly, the SNR of each hybridised spot was determined by:

$$SNR = \frac{\left(Intensity_{total} - Intensity_{background}\right)}{Intensity_{background}}$$

Spots with an SNR below a threshold value were considered not significantly different from background and were eliminated from analysis. This was a conservative approach and may remove some low intensity but authentic spots. SNR thresholds of 2.0 for Cy5 and 1.5 for Cy3 have been previously suggested to eliminate noisy data while retaining majority of 'good' data (Jenner et al., 2001). The determination of the SNR threshold value for data filtering in HHV-7–human microarrays was performed as follows.

Three microarrays (S02, S29, S30) were hybridized to Cy5-labelled RNA from mock-infected cells and Cy3-labelled reference RNA. Since HHV-7 RNA was absent in the mock-infected cells, any Cy5 fluorescence intensities detected by the HHV-7 probes were due to noise or unspecific hybridisation. The mean Cy5 intensity (of nine repeats – three slides with three spots per slide) of the 96 HHV-7 spots ranged between 60 and 250 fluorescence units (Figure 3.10a). In order to select the SNR threshold which achieves the maximum removal of noisy data while retaining as much 'good' data as possible, the number of probes that passed the SNR filtering in both channels were plotted against increasing values of SNR threshold (Figure 3.10b). The number of false positives in the Cy5 channel decreased rapidly with increasing SNR threshold, while the number of 'good' data points in the Cy3 channel showed a much slower drop. At the SNR of 1.5, all 6 negative control blank spots were removed in both Cy5 and Cy3 channels, hence was regarded as the lowest acceptable value for SNR filtering purposes. However, the number of noisy data points in the Cy5 channel was still high (44 of 90 non-blank spots, or 49.9%). The SNR threshold of 2.0 was the highest value at which

the pass-rate in the Cy3 channel remained above 90% (85 out of 90, or 94%), while only 8 of 90 spots (8.9%) in the Cy5 channel were higher than the threshold (DR6, DR6', U7, U19, U36, U42, X5 and U79) (see specificity section 3.2.4.4). Therefore by raising the threshold from 1.5 to 2.0, the number of noisy data points in Cy5 was reduced by more than 5-fold, while the number of data points in Cy3 was only reduced by 4%, causing 3 additional spots (U49, U71, one blank) to fail the filtering.

The software used for data filtering (section 2.9) did not allow different threshold values to be set for the two channels, hence it was not possible to filter Cy5 and Cy3 data separately with respective SNR thresholds of 2.0 and 1.5 as described previously (Jenner et al., 2001). When SNR threshold of 2.0 was used instead of 1.5, the pass-rate in the Cy3 channel was reduced by 4%, but an extra 36 false positives in the Cy5 channel were removed. Therefore a SNR threshold of 2.0 was selected for filtering the HHV-7 microarray data, to obtain a lower false positive rate.

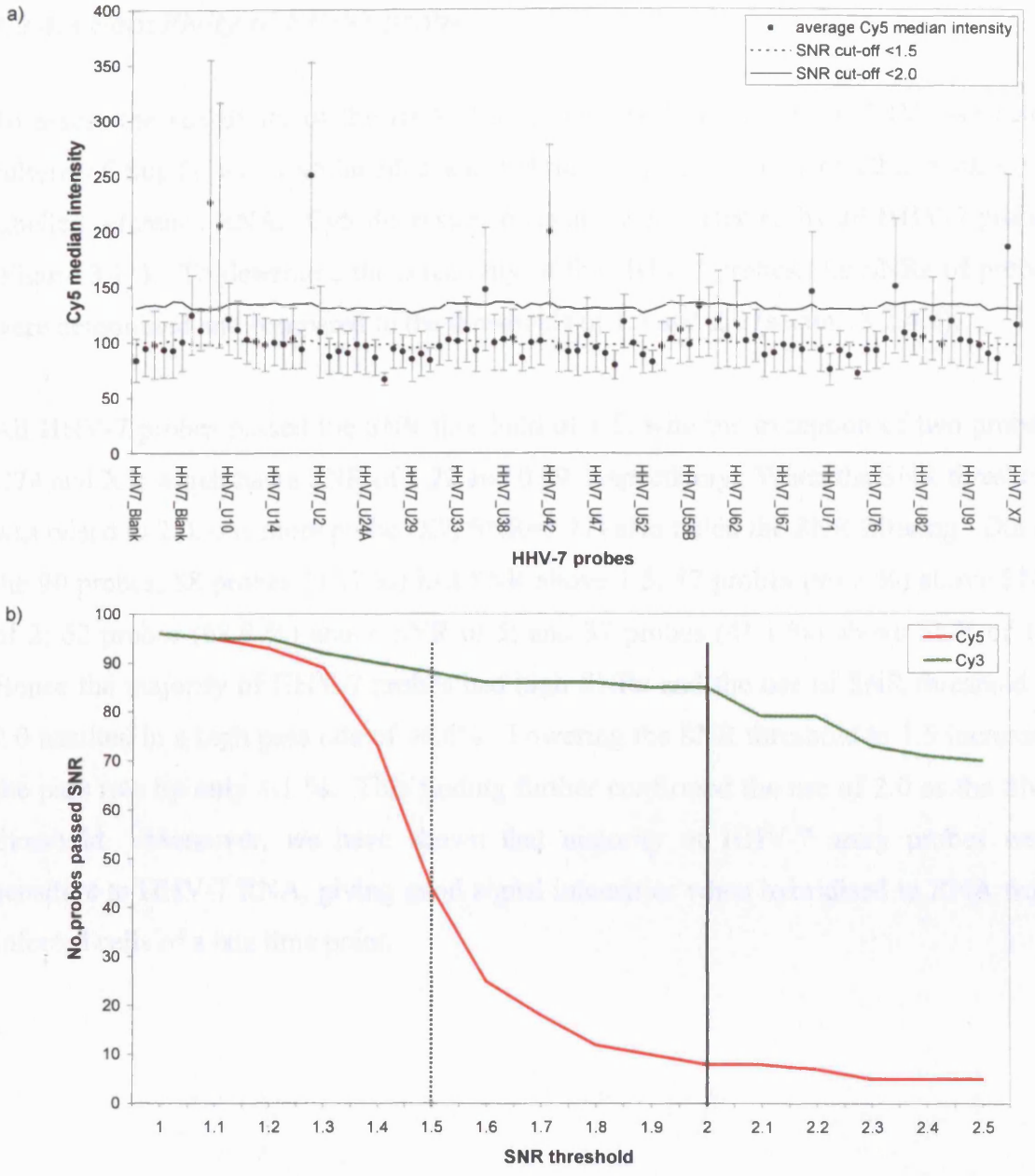


Figure 3.10 Three microarrays (S02, S29 and S30) were hybridised to Cy5-labelled RNA from mock-infected cells and Cy3-labelled reference RNA. a) Most of the Cy5 fluorescence intensities (average of 9 replicates, error bars represent standard deviation) detected by the HHV-7 probes (black dots) were below the SNR threshold of 2.0 (solid line); some of these were above the SNR threshold of 1.5 (dotted line). b) The number of probes passing SNR filtering decreased with increasing SNR threshold in both Cy5 (red) and Cy3 (green) channels. An SNR threshold of 1.5 or above (dotted line) was required to filter out all 6 blank spots. The SNR threshold of 2.0 (solid line) was selected to filter out maximum number of probes in Cy5 while retaining maximum number of probes in Cy3.

3.2.4.3 Sensitivity of HHV-7 probes

To assess the sensitivity of the HHV-7 array probes, RNA from an HHV-7–infected culture of SupT1 was Cy5-labelled and hybridised to slide S17, together with Cy3-labelled reference RNA. Cy5 fluorescence signals were detected by all HHV-7 probes (Figure 3.11). To determine the sensitivity of the HHV-7 probes, the SNRs of probes were determined and compared to the thresholds of 1.5 and 2.0 (section 3.2.4.2).

All HHV-7 probes passed the SNR threshold of 1.5, with the exception of two probes, U74 and X3, which had a SNR of 1.22 and 0.99, respectively. When the SNR threshold was raised to 2.0, one more probe (X7, SNR=1.71) also failed the SNR filtering. Out of the 90 probes, 88 probes (97.7 %) had SNR above 1.5; 87 probes (96.6 %) above SNR of 2; 62 probes (68.9 %) above SNR of 5; and 37 probes (41.1 %) above SNR of 10. Hence the majority of HHV-7 probes had high SNRs and the use of SNR threshold at 2.0 resulted in a high pass rate of 96.6%. Lowering the SNR threshold to 1.5 increased the pass rate by only 1.1 %. This finding further confirmed the use of 2.0 as the SNR threshold. Moreover, we have shown that majority of HHV-7 array probes were sensitive to HHV-7 RNA, giving good signal intensities when hybridised to RNA from infected cells of a late time point.

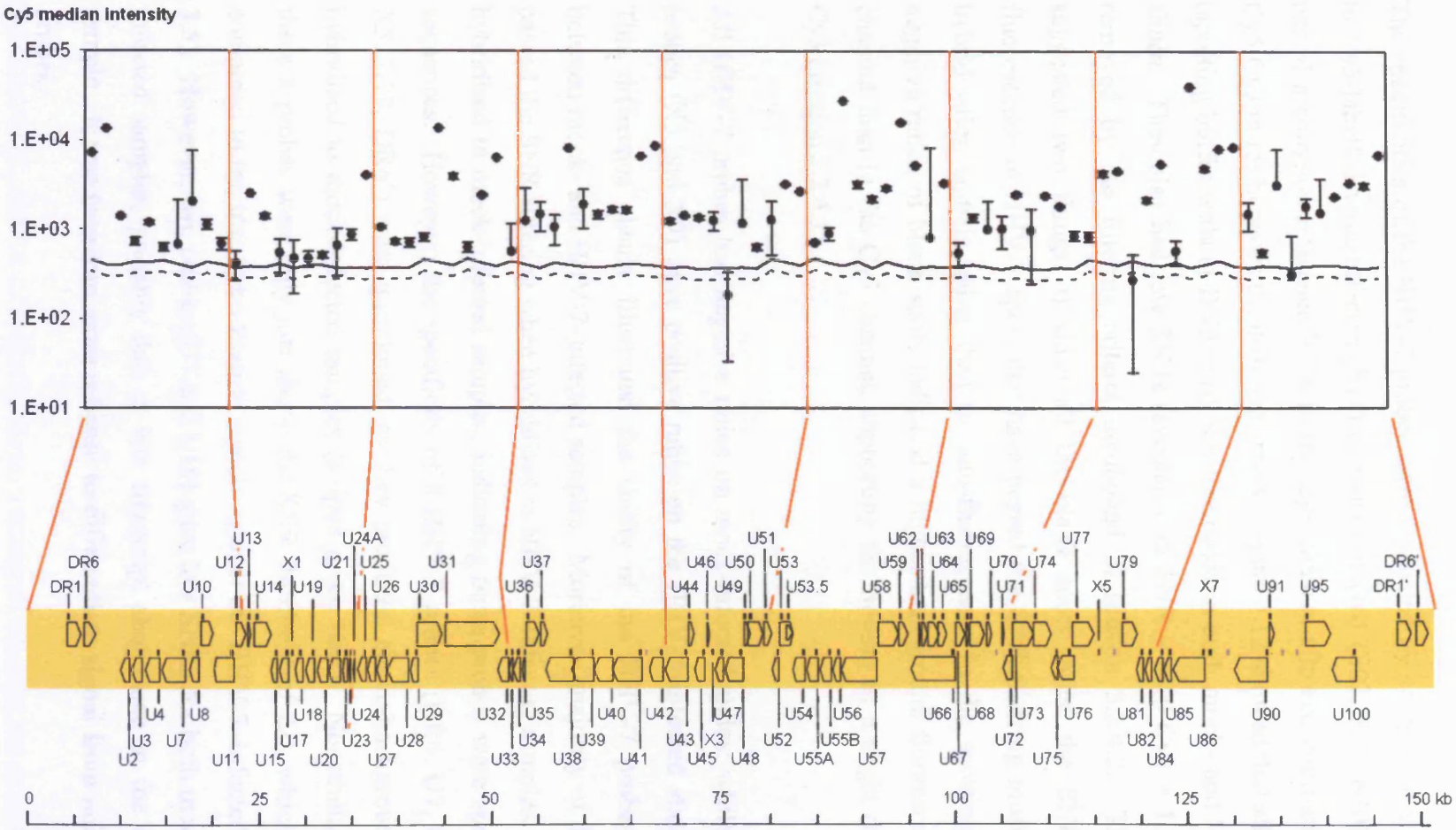


Figure 3.11 Microarray S34 was hybridised to Cy5-labelled RNA from HHV-7 infected cells (72 hpi) and Cy3-labelled reference RNA. Average fluorescence signals were calculated of triplicate spots (black dots). All except two probes (U74, X3) were above the signal-to-noise threshold of 1.5 (dotted line). Raising the SNR threshold to 2.0 (solid line) resulted in probe X7 also failing SNR filtering. Gene probes on graph in the same order as the corresponding ORFs in genome. 90 array probes were successfully amplified (black rectangles) and 6 did not amplify (grey rectangles). Guide lines (orange) are included to aid tracking the data points to their corresponding ORFs.

3.2.4.4 Specificity of microarrays

The specificities of the HHV-7 probes were evaluated by comparing arrays hybridised to Cy5-labelled virus-infected (S17) or mock-infected (S02, S29, S30) samples. The use of a common reference RNA in the Cy3 channel allowed comparison of the $\log_2$ Cy5/Cy3 ratios between the different arrays. Figure 3.12 showed that all six blank spots (spotting buffer without DNA) had negative ratios in both mock- and HHV-7–infected slides. They also had low SNRs (average Cy5 SNR = 1.40 Cy3 = 1.39) hence were removed by the filtering criteria mentioned in section 3.2.4.2. This observation suggested two things: i) since all the blank spots failed the SNR filtering, the fluorescence in HHV-7 spots that have passed the SNR filtering must be due to real hybridisation signals rather than to auto-fluorescence in the spotting buffer; ii) the negative ratios of blank spots indicated a higher background fluorescence in the Cy3 channel than in the Cy5 channel, supporting the finding of a slight dye-bias towards Cy3 (section 3.2.4.1).

All HHV-7 probes had negative ratios on mock-infected slides; while all except two probes (X3 and X7) gave positive ratios on the HHV-7–infected slide (Figure 3.12). This difference clearly illustrated the ability of the HHV-7 probes to distinguish between mock- and HHV-7–infected samples. Moreover, majority of the HHV-7 spots passed the SNR filtering when hybridised to HHV-7–infected samples, but failed when hybridised to mock-infected samples, indicating these probes were specific to HHV-7 sequences. However, the specificity of 8 HHV-7 probes (DR6, U7, U19, U36, U42, X5, U79, DR6') were questioned, as they produced above-background signals when hybridised to mock-infected samples (striped green bars). Nonetheless, the SNRs of these 8 probes were only just above the SNR threshold of 2.0, which were very low compared to the true hybridisation signals against the HHV-7–infected samples (Table 3.5). However, two probes (U7 and U19) gave low SNRs for both mock- and HHV-7–infected samples, possibly due to low transcript abundance in the HHV-7–infected sample. It was therefore more difficult to differentiate signal from noise for these two probes.

The sequences of these 8 probes gave no significant matches to human genes by BLAST searches. In fact, the probes only produced significant matches to their target positions in HHV-7 genome (Table 3.2), except U36 and U42, which also blasted to the corresponding conserved sequences in HHV6A and HHV6B, with lower similarities (62% and 66% similarity respectively) and lower E-values (10^{-9} and 10^{-13} respectively). Therefore the reason for these 8 probes passing the SNR filtering was more likely to be due to a non-perfect filtering method (i.e. false positives) rather than cross-hybridisation with cellular transcripts in the uninfected cells. These data showed the HHV-7 array had a high degree of specificity to HHV-7 RNA. Although all probes were analysed in future experiments, 11 of the probes (DR6, DR6', U7, U19, U36, U42, U74, U79, X3, X5 and X7) were considered with caution due to either high SNR in mock infected cells (section 3.2.4.2) or low SNR in sensitivity test (section 3.2.4.3).

Table 3.5 Comparison of SNR for a subset of HHV-7 probes

Probe	SNR (mock-infected)	SNR (HHV-7–infected)
DR6	3.49	67.65
U7	2.18	2.81
U19	3.70	2.61
U36	2.21	40.74
U42	2.95	44.38
X5	2.85	212.64
U79	2.25	23.89
DR6'	3.14	35.66

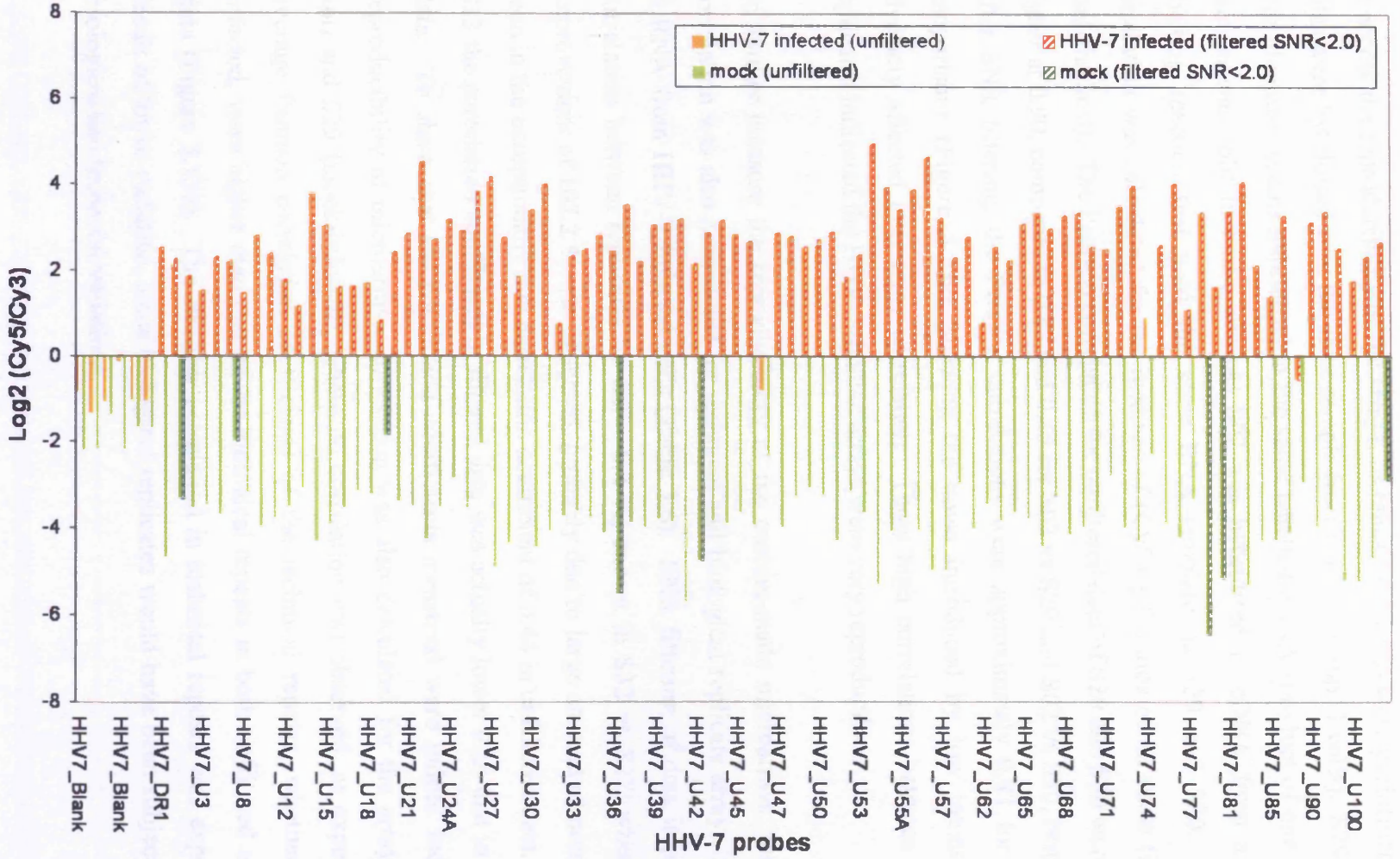


Figure 3.12 Microarrays were hybridised to Cy5-labelled RNA from HHV-7 infected cells (red bars) or mock infected cells (green bars) and Cy3-labelled reference RNA. The average log2 Cy5/Cy3 ratios were calculated for data before (solid bars) and after (striped bars) signal-to-noise ratio filtering.

3.2.4.5 Reproducibility of microarrays

To access the reproducibility of the arrays, technical replicates were performed. Three slides were hybridised with the same sample (mock-infected SupT1 cells). S29 and S30 were identical repeats hybridised to the same labelled cDNA (product of one labelling reaction was split in half); whereas S02 was hybridised to cDNA from a separate labelling reaction that used the same RNA template as S29 and S30. Pearson correlation was calculated for the medians of (Cy5/Cy3) ratios of all spots (including viral and host). The Pearson correlation for unfiltered data of S29 and S30 were slightly higher at 0.90, compared to 0.88 and 0.85 for S02 vs S29 and S02 vs S30, respectively. After SNR filtering, the Pearson correlations were approximately 0.93 for all three comparisons (Figure 3.13a), showing the noise introduced by low intensity spots adversely affected inter-array correlation. These high correlations between technical replicates indicated the HHV-7–human arrays were very reproducible.

To further measure the reproducibility of the custom-made microarrays, the Pearson correlation was also determined for technical and biological replicate arrays hybridised to RNA from HHV-7 infected cells (Table 3.6). SNR filtering of data improved the correlation between replicates by 0.8 – 8.8 %, except in S32 vs S27, where a large improvement of 102.2 % was observed, probably due to large amount of noisy data, as seen in the exceptionally low correlation coefficient of 0.44 in unfiltered data. In S33 vs S13 the correlation coefficient in filtered data was actually lower than that in unfiltered data. To show that the correlation coefficients measured were indeed indicative of reproducibility of microarrays, correlation was also calculated for the arrays S34 (72 hpi) and S29 (mock-infected), where no correlation was observed, as expected. The average Pearson correlation coefficients of the technical repeats, whether mock or infected, were higher than that of the biological repeats in both unfiltered and filtered data (Figure 3.13b). The stronger correlation in technical repeats was expected, as a result of lower variation, since biological replicates would have been subjected to both biological and technical variations.

Table 3.6 Reproducibility of microarrays – Pearson correlation coefficients

		Pearson correlation coefficient		
		Unfiltered	Filtered	Improve by
Mock technical		**Ave 0.88**	**Ave 0.93**	
S02 Mock	S29 Mock	0.88	0.93	5.6%
S02 Mock	S30 Mock	0.85	0.93	8.8%
S29 Mock	S30 Mock	0.90	0.93	3.5%
Infected technical		**Ave 0.92**	**Ave 0.95**	
S33 48hpi	S19 48hpi	0.96	0.97	0.8%
S34 72hpi	S17 72hpi	0.74	0.73	-1.6%
Infected biological		**Ave 0.71**	**Ave 0.83**	
S33 48hpi	S13 48hpi	0.87	0.94	7.2%
S34 72hpi	S28 72hpi	0.90	0.91	1.3%
S15 48hpi PAA	S11 48hpi PAA	0.74	0.79	7.2%
S32 24hpi	S27 24hpi	0.44	0.89	102.2%
Infected vs Mock				
S34 72hpi	S29 Mock	-0.05	0.12	N/A

In summary, the HHV-7–human microarrays produced highly correlated data amongst technical replicates (average correlation after filtering = 0.93 – 0.95) and biological replicates (average correlation post-filter = 0.83). These values were comparable with previously published data for both technical replicates: 0.96 (DeRisi et al., 1997), 0.91 (Ahn et al., 2002), 0.87 (Cohen 2000), 0.84 - 0.86 (Mayne et al., 2001), and biological replicates 0.97 (Scherf et al., 2000), 0.90 (Zhang et al., 2001b), 0.83 - 0.92 (Ross et al., 2000). Therefore the reproducibility of the HHV-7–human microarrays were satisfactory.

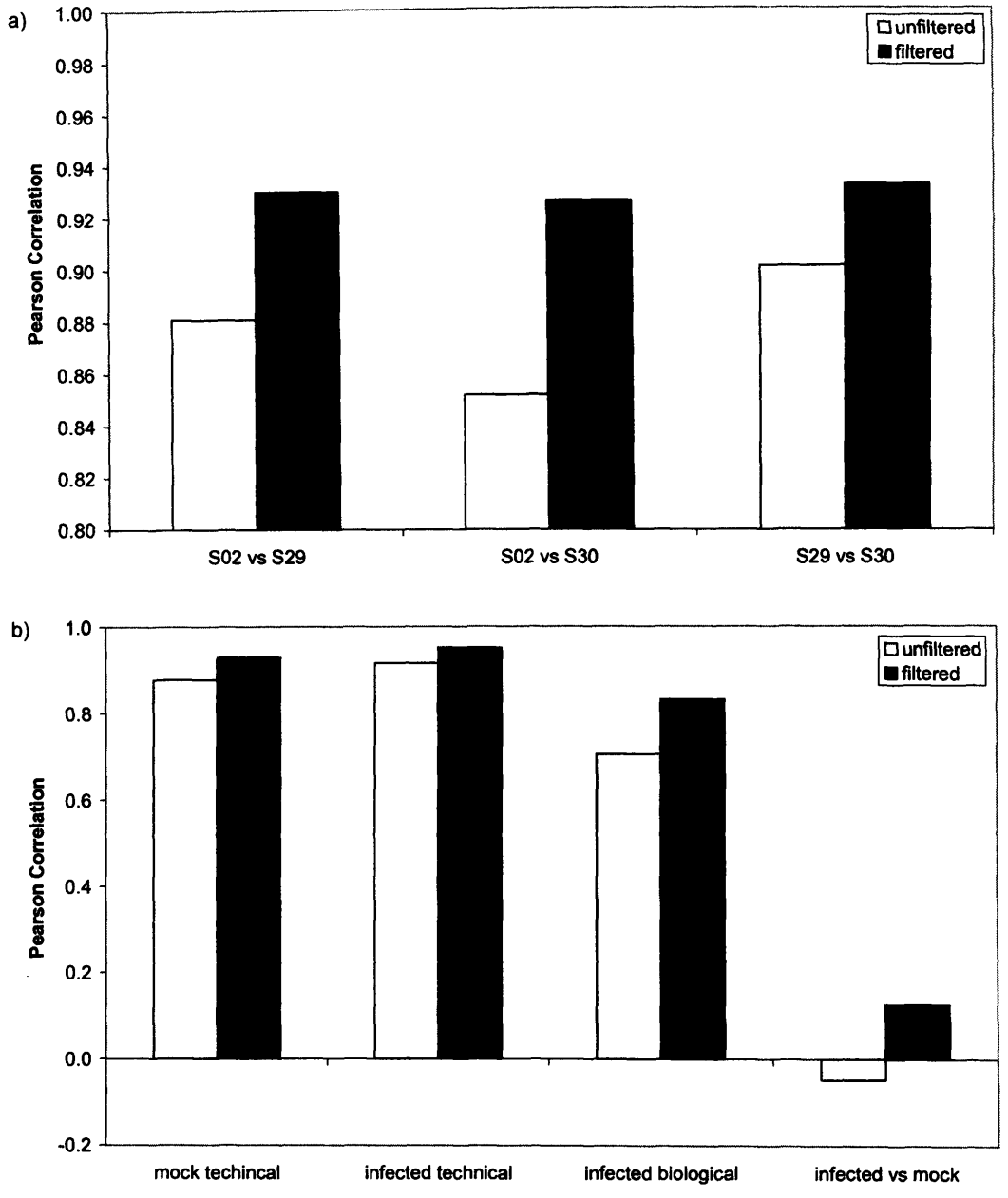


Figure 3.13 Pearson correlation coefficients of median (Cy5/Cy3) ratios between replicate slides before (white bars) and after (shaded bars) SNR filtering. a) Microarrays S29 and S30 were identical technical replicates hybridised to the same labelled cDNA in both Cy5 (mock-infected) and Cy3 (reference RNA); S02 was hybridised to cDNA of a different labelling reaction using the same RNA as S29 and S30. b) Pearson correlation coefficients of technical repeats of mock-infected (S02/S29/S30); technical repeats of HHV-7 infected (S33/S19, S34/S17); biologcial repeats of HHV-7 infected (S33/S13, S34/S28, S15/S11, S32/S27); and infected vs mock. Also see Table 3.6

3.2.4.6 Use of HHV-7–human microarray in gene expression profiling

The usability of the HHV-7–human microarrays in gene expression profiling of HHV-7–infection was tested. A selection of microarray data were analysed by clustering (section 2.10) to test whether the arrays were able to produce distinguishing patterns for different samples. Arrays were arranged into nodes in a tree according to the similarity of their patterns (Figure 3.14). Several findings were observed. Firstly, replicates were clustered together, indicating that similar expression patterns were detected by the arrays. Furthermore, technical replicates were more similar than biological replicates (S33 and S19; S07 and S13), consistent with results obtained in section 3.2.4.5. The arrays were even able to recognise identical repeats (S29 and S30) as more similar than non-identical ones (S02). Perhaps the most important observation was that the arrays were able to produce distinguishing patterns not only between mock- and HHV-7–infected samples, but also between samples at different time point post-infection (8 hpi and 48 hpi). It can be seen that the variation between different time points was higher than the variation between biological replicates of a particular time point.

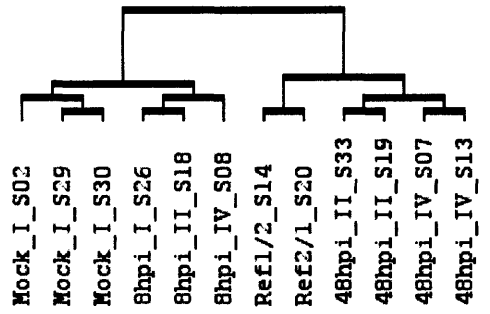


Figure 3.14 Tree structure of clustering analysis of selected microarrays.

Arrays with most similar patterns are clustered together. Samples with the same Roman numbers are technical replicates; different Roman numbers denote biological replicates

3.2.5 Data Normalisation

In order to ensure accurate measurement of biological-based differential gene expression in HHV-7 infected cells, microarray data were normalised (after filtering) to remove experimentally introduced variations, including different efficiencies in labelling, hybridisation, and effects introduced during scanning. Microarray data were subjected, in sequential order, to total intensity normalisation and LOWESS normalization using the Microarray Data Analysis System (MIDAS). MIDAS is part of a suite of programs known as TM4, developed at The Institute for Genomic Research (TIGR) (Saeed et al., 2003).

In this section, data from arrays S14 and S20 were shown as examples. Since these arrays were hybridised to reference RNA in both Cy5 and Cy3 channels, the signal in both channels should be equal, meaning the effectiveness of each normalisation step could be assessed by comparing real data to an idealised view of the data.

3.2.5.1 Effects of Total Intensity Normalisation

Due to differences in labelling efficiencies, most arrays in this study were systematically biased towards the Cy3 channel (section 3.2.4.1). Array S14 hybridised to reference RNA in both Cy5 and Cy3 channels showed a high Pearson correlation coefficient between the two dye channels (section 3.2.4.1) but also had a Cy3-bias, as illustrated by the negative $\log_2$(Cy5/Cy3) ratios (Figure 3.15a); and this bias was not limited to a specific area of the slide, but a common phenomenon across the whole slide (Figure 3.15b). In order to correct for this dye-bias, data were adjusted by a normalisation factor (section 2.9) so that the total intensities of all spots were equal in both channels. This resulted in an upward shift, where the $\log_2$(Cy5/Cy3) ratios of all the spots were increased proportionally to their original values (Figure 3.15c, 3.15d). Data from S20 were also subjected to total intensity normalisation, but the effect was not as noticeable because the two channels were already fairly balanced in raw data (Figure 3.16a-d).

3.2.5.2 Effects of Lowess Normalisation

Total intensity normalisation assumed that the variation in the $\log_2$(Cy5/Cy3) ratios of the data were independent of signal intensity. In reality, lower intensity spots tend to have $\log_2$(Cy5/Cy3) ratios that deviate more from zero (section 1.2.3.2), as shown in the raw data of array S20 (Figure 3.16a). This phenomenon could not be corrected by the use of a common normalisation factor in total intensity normalisation (Figure 3.16c).

Local weighted scatterplot smoothing (Lowess) normalisation was used to correct for intensity-dependent variation by applying a local weighted linear regression as a function of the intensity, with the Cy3 as reference point, as previously suggested (Quackenbush, 2002). The Lowess-adjusted data had a more balanced distribution of median $\log_2$(Cy5/Cy3) ratios vs intensity (Figure 3.16e). The number of spots with $\log_2$(Cy5/Cy3) ratios below -1.0 was 9 before normalisation and 0 after normalisation; For above +1.0, these numbers were 2 and 7, respectively. Therefore intensity-dependent variation was reduced by Lowess normalisation, preparing the data for more accurate analyses downstream.

Although one may argue that Lowess normalisation is an artificial adjustment of data, and may result in loss of definition in low abundance transcripts, it allows the correction of data points with low signal and falsely high ratios which may otherwise appear as highly differentially expressed genes. A previous study has shown that although the spread of data was reduced by Lowess normalisation, the ability to detect genes known to be differentially expressed was not affected (Yang et al., 2002b).

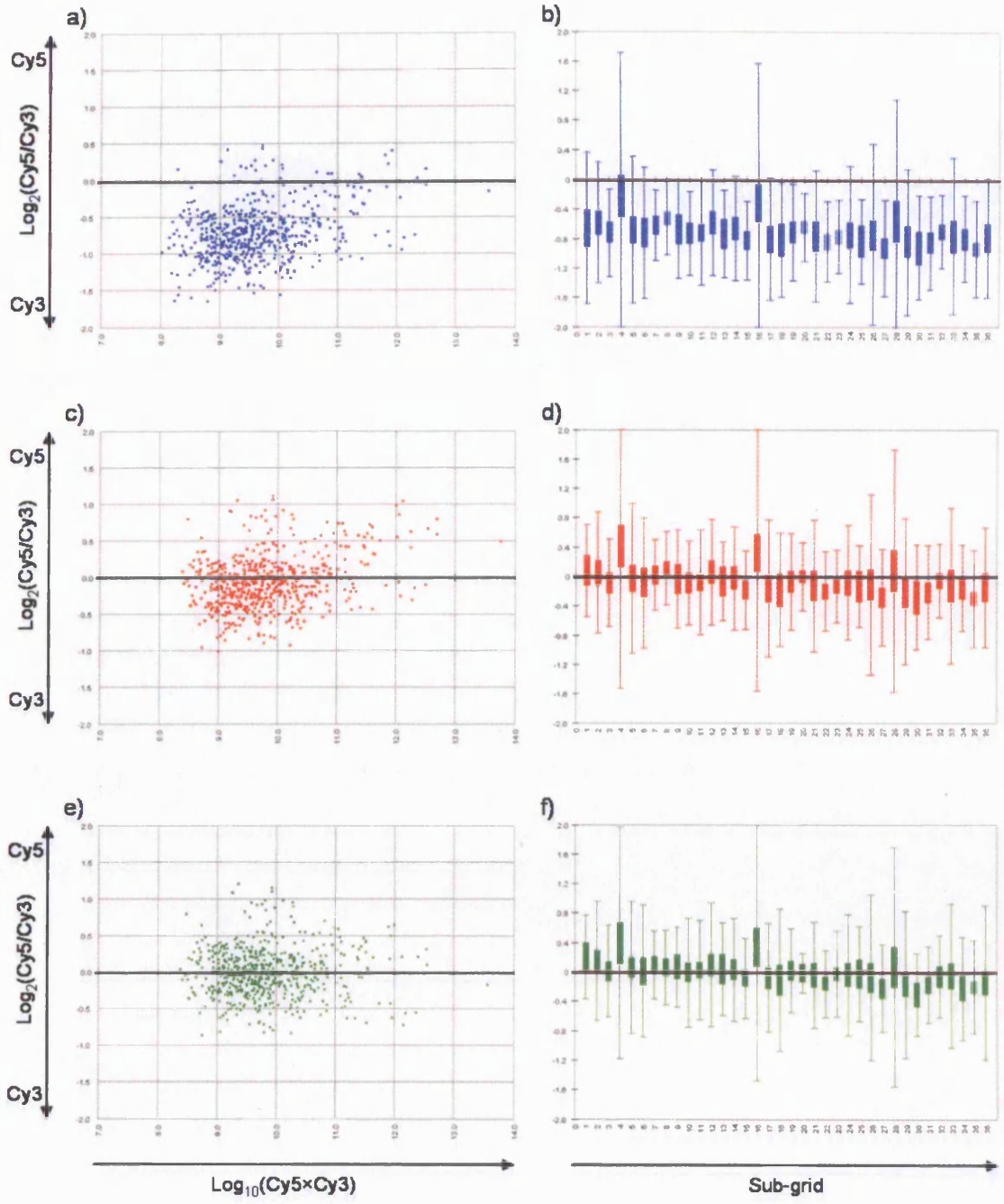


Figure 3.15 Microarray data of S14, hybridised to reference RNA labelled in Cy5 and Cy3, shown in M/A plots (a,c,e) and box plots (b,d,f). Raw data (a-b) were adjusted by total intensity normalisation (c-d); followed by Lowess normalisation (e-f).

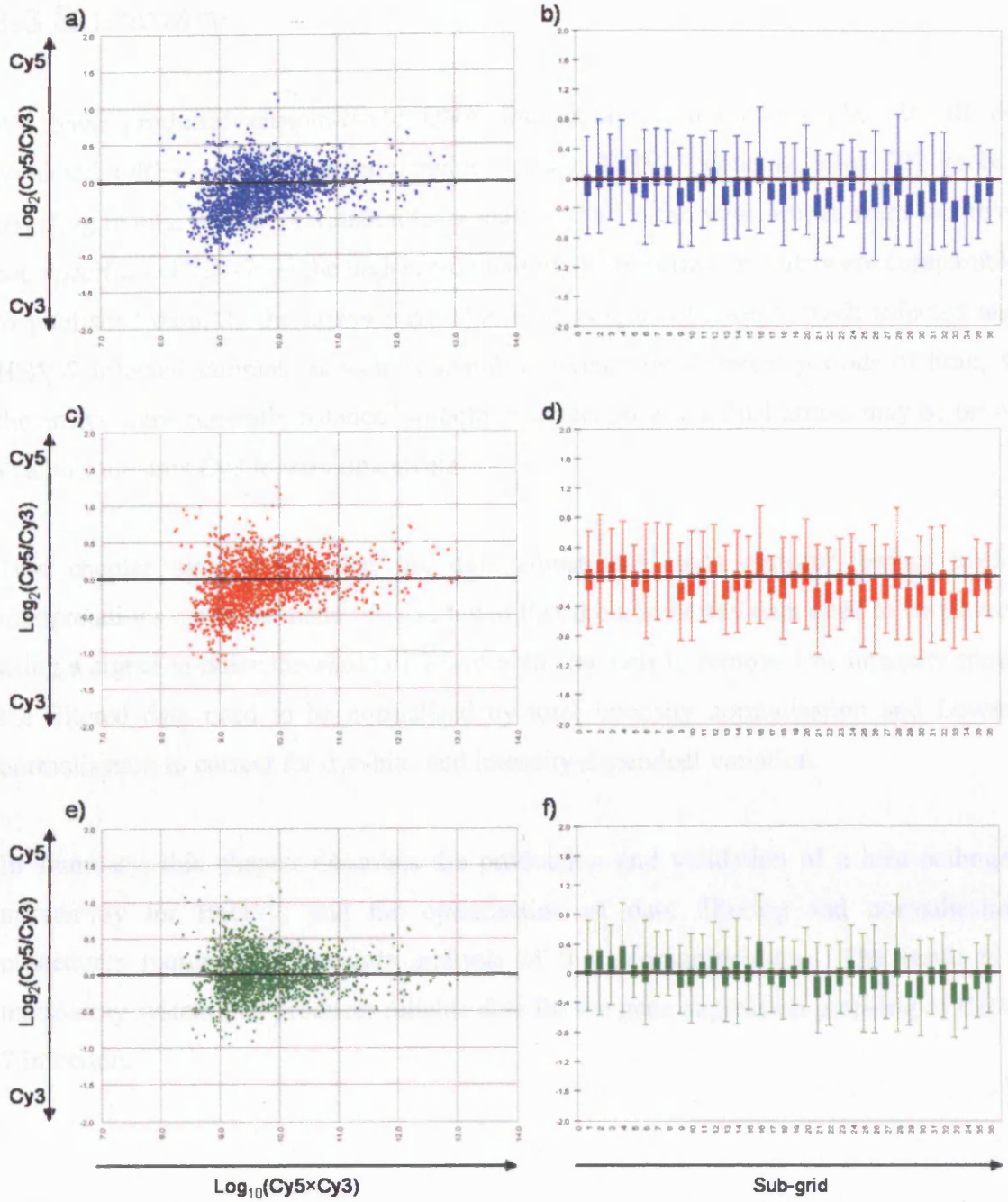


Figure 3.16 Microarray data of S20, hybridised to reference RNA labelled in Cy5 and Cy3, shown in M/A plots (a,c,e) and box plots (b,d,f). Raw data (a-b) were adjusted by total intensity normalisation (c-d); followed by Lowess normalisation (e-f).

3.3 Summary

We have produced custom-made DNA microarrays containing probes to all 86 predicted ORFs and 4 intergenic regions of the HHV-7 genome, plus 907 probes targeting human genes. Validation tests showed that 1) the viral probes were sensitive and specific to HHV-7; 2) the high reproducibility of the arrays results were comparable to published data; 3) the arrays were able to distinguish between mock infected and HHV-7 infected samples, as well as samples infected for different periods of time; 4) the arrays were generally balanced in both channels but individual arrays may be prone to a bias towards Cy3 to varying extents.

This chapter also investigated the data adjustment steps required before useful interpretations could be made. It was found that the microarray data need to be filtered using a signal-to-noise threshold of 2.0 in both channels to remove low intensity spots; the filtered data need to be normalised by total intensity normalisation and Lowess normalisation to correct for dye-bias and intensity-dependent variation.

In summary, this chapter describes the production and validation of a host-pathogen microarray for HHV-7; and the optimisation of data filtering and normalisation procedures required for accurate analysis of the microarray data. The result is a microarray system that produces reliable data for the gene expression profiling of HHV-7 infection.

CHAPTER 4

INFECTIVITY OF HUMAN HERPESVIRUS TYPE-7 (HHV-7) IN SUPT1 CELLS

4.1 Introduction

In order to study the expression patterns of viral and host genes during human herpesvirus type-7 (HHV-7) infection, a time course infection of SupT1 cells by cell-free HHV-7 virus was required, and the transcription profiles of the infected cells analysed using the custom-made HHV-7 microarrays. To generate sufficient mRNA (200-500 ng) for each microarray analysis, 2×10^6 cells per infection are needed. Ideally a multiplicity of infection (MOI) of 1.0 or above is desirable to ensure majority (67 %) of the cells are infected, so that the expression patterns obtained are representative of HHV-7 infected cells rather than uninfected ones. Therefore this chapter is concerned with developing and optimising procedures for growing the HHV-7 virus in cell culture, producing high titre virus stocks, accurately quantifying the virus stocks, infecting SupT1 cells at a MOI of 1.0 or above, and using RNA amplification to facilitate microarray analysis.

4.1.1 Infectivity of HHV-7

HHV-7, like other betaherpesviruses (Shiraki et al., 1991), has long been known to be a low titre virus (Pellett, 2001), and have low infectivity-to-viral particle ratios (less than 1 infectious unit per 10^4 particles) (Ablashi et al., 1998), which has hindered gene expression studies of the virus (Pellett, 2001). Such low ratios have been observed in human cytomegalovirus (HCMV), and are attributed to the production of noninfectious enveloped particles and dense bodies (Stinski et al., 1979), but it is unknown whether HHV-7 produce such particles. Direct pelleting by ultra-centrifugation resulted in higher viral particles/ml but not higher infectivity titre (Ablashi et al., 1998), probably because the virus was damaged during ultra-centrifugation. HHV-7 is a cell-associated virus, and infected cell lysates have been shown to be about 10 times more infectious than virus supernatant (Ablashi et al., 1998). However, in order to accurately monitor the changes in gene expression over time, a homogenous infection is required in this study. Therefore it is necessary to use cell-free virus and to remove the virus after the inoculation period. This constraint has limited the infectious titre used in this study. In fact, the low infectivity of HHV-7 has limited past attempts on gene expression studies

(Pellett, 2001), therefore highlighting the importance of producing high titre stocks in this project.

4.1.2 Spinoculation and Polybrene

Different methods have been developed to improve the infectivity of viruses with low infectivity-to-particle ratios, including the use of spinoculation (also known as centrifugal infection) amd the addition of cationic polymers such as polybrene. Spinoculation has been applied to successfully increase the infectivity of many viruses, for example, HIV-1 (Ho et al., 1993; O'Doherty et al., 2000; Pietroboni et al., 1989), herpes simplex virus (Tenser and Dunstan, 1980), human herpesvirus type-8 (HHV-8) (Inoue et al., 2003), cytomegaloviruses (Ho et al., 1993; Hodgkin et al., 1988; Hudson, 1988; Hudson et al., 1976), HHV-6 (Pietroboni et al., 1989), and human adenovirus (Nyberg-Hoffman et al., 1997). The enhancement effect of spinoculation on viral infectivity has been attributed to several factors that are believed to be increased by centrifugation, including cellular receptivity (Hudson, 1988; Hudson et al., 1976), viral fusion (Tenser and Dunstan, 1980), and viral deposition (Hodgkin et al., 1988; O'Doherty et al., 2000; Osborn and Walker, 1968), bringing viruses to closer contact with the cells. A study on influenza A virus had shown that reducing the volume of virus inoculum has the same effect as centrifugation (Reading et al., 2001) therefore favouring the close-contact hypothesis. The presence of polybrene (2-8 μg/ml) has been used to enhance the infectivity of HHV-8 (Inoue et al., 2003) and HIV (Pan et al., 1993), but had no effect on HHV-6 infectivity (Black et al., 1989). The addition of polybrene, together with an anionic polymer, has also been shown to enhance retrovirus infection of cells in a concentration-dependent manner (Landazuri and Le Doux, 2004), due to formation of virus–polymer complexes which aided sedimentation and delivered the viruses to the cells more rapidly than by simple diffusion.

4.1.3 RNA amplification

The application of the RNA amplification technique in gene expression analysis has been well documented (McClintick et al., 2003; Petalidis et al., 2003; Rajeevan et al.,

2003; Schindler et al., 2005), especially for detecting low copy number gene transcripts (Van Gelder et al., 1990), and for RNA extracted from limiting amounts of starting materials obtained by biopsy, laser capture microdissection (Luo et al., 1999) and flow cytometry-sorting. A number of different methods for RNA amplification have been published, which can be roughly divided into polymerase chain reaction (PCR)-based (Frohman et al., 1988) and *in vitro* transcription (IVT)-based approaches. Both methods essentially involve cDNA synthesis from the RNA by reverse transcription, followed by amplification of the cDNA. PCR based methods amplify the cDNAs using *Taq* polymerase. Due to the low fidelity of the enzyme and the exponential nature of the reaction, PCR-based methods have been largely replaced by the IVT-based approach, also known as the Eberwine method (Van Gelder et al., 1990). This approach involves the incorporation of a T7 promoter sequence in the synthesis of first strand cDNA using a specific cDNA synthesis primer, followed by second strand synthesis by DNA polymerase I. Antisense RNAs are then generated by IVT using T7 RNA polymerase. This method has the advantage over PCR-based approaches in that RNA is amplified in a linear rather than exponential manner. However, the aRNAs generated by the Eberwine method are anti-sense to the mRNA and 3'-biased, due to the use of oligo-dT primer in first strand synthesis (Nam et al., 2002). Therefore recent methods have made modifications to correct for this bias, such as the use of random primers (Goff et al., 2004), template switching (Petalidis et al., 2003; Puskas et al., 2002; Seth et al., 2003), and terminal continuation (Che and Ginsberg, 2004). Methods have also been developed to amplify sense strand RNA that are in the same orientation as the mRNA transcript (Goff et al., 2004; Marko et al., 2005; Rajeevan et al., 2003), which are suitable for use with oligonucleotide arrays and for array probes that require RNAs without a 3'-bias. RNA amplification has been shown to generate results that were reproducible within each method, although minor, method-specific systematic biases have been detected (Schindler et al., 2005; Wilson et al., 2004) when compared with unamplified RNA. Nonetheless, as long as the bias is consistent in all samples, such effects may be cancelled out in two-colour spotted microarrays, which measure relative expression ratios between two samples (Goff et al., 2004). Moreover, it has been shown that the correlation between aRNAs amplified by the T7-based method was good and the overall bias was low (Zhao et al., 2002), and the template switching method was able to preserve the known relative abundance of plant mRNA spiked into human mRNA samples at proportions between 0.1% to 0.00001% (Rajeevan et al., 2003).

RNA amplified using the SenseAmp kit (Genisphere, USA) has been shown to be reproducible, as well as having a high Pearson correlation coefficient of 0.90 when compared to unamplified RNA using quantitative real-time PCR (Goff et al., 2004).

4.1.4 Scope of chapter 4

The aim of this chapter was to establish a high MOI in SupT1 cells with HHV-7 cell-free virus, and to produce sufficient RNA for subsequent gene expression analysis. For this purpose the suitability of spinoculation, polybrene, and RNA amplification for HHV-7 infection of SupT1 cells was explored.

4.2 Results

4.2.1 Maintaining HHV-7 infection in SupT1 cells

HHV-7 (strain DC) infection of SupT1 cells *in vitro* was established in the laboratory. Infection was maintained by co-culture at a ratio of 1:1 to 1:2 (infected-to-uninfected cells). This culture ratio had to be carefully followed, because over-dilution with uninfected cells resulted in the loss of active infection. This was probably because HHV-7 has a slow infection profile hence the proliferating uninfected cells quickly outgrew the spread of infection. This finding agreed with the observation that HHV-7 grew best in media supplemented with a lower concentration (5%) of FBS (normally 10% is used for SupT1), which retarded the proliferation of uninfected cells (Ablashi et al., 1998). However, our results were contrary to the 1:5 co-culture ratio previously suggested by Secchiero *et al* (Secchiero et al., 1997a), who used a different HHV-7 isolate (AL), which may have different growth kinetics. Some isolates have indeed been shown to be more cytopathic than others in SupT1 cells (Ablashi et al., 1998). The infection kinetics of the DC strain had not been determined in this study.

4.2.2 Morphology of infected cells

Infected cultures displayed cytopathic effect (CPE), characterised by the formation of giant multi-nucleated syncytia and attached cells – large, elongated, flattened cells that were loosely attached to the base of the culture container (Figure 4.1), consistent with previous studies (Ablashi et al., 1998). Flow cytometry analysis (FACS) of infected cultures revealed a population of larger and more granular cells, indicated by higher forward scatter (FSC) and side scatter (SSC) values, respectively (Figure 4.2e, h), in contrast to the small uninfected cells (Figure 4.2b). When comparing a cell-free virus-infected culture 4 days post inoculation to an ongoing infection maintained by co-culture, the degree of CPE (proportion of large and granular cells) did not correlate with the proportion of cells bearing HHV-7 antigen as detected by FACS and immunofluorescent microscopy (Figure 4.2). The percentage of granular cells exceeded those bearing HHV-7 antigen. It may be possible that these cells displaying CPE were not infected, but experienced morphological changes as bystander effect. A similar cellular phenotype had been observed previously, where apparently uninfected cells in an infected culture underwent apoptosis (Secchiero et al., 1997a). As a result, visual judgement of CPE may be falsely high in the presence of cell debris in an ongoing culture, whereas a culture infected by cell-free virus may contain a portion of infected, HHV-7 antigen bearing cells that have no visible CPE. In summary, these finding demonstrated that although CPE may be crudely indicative of the degree of infection, it might not give accurate results as a direct measure of virus titre. This will be discussed further in the following sections.

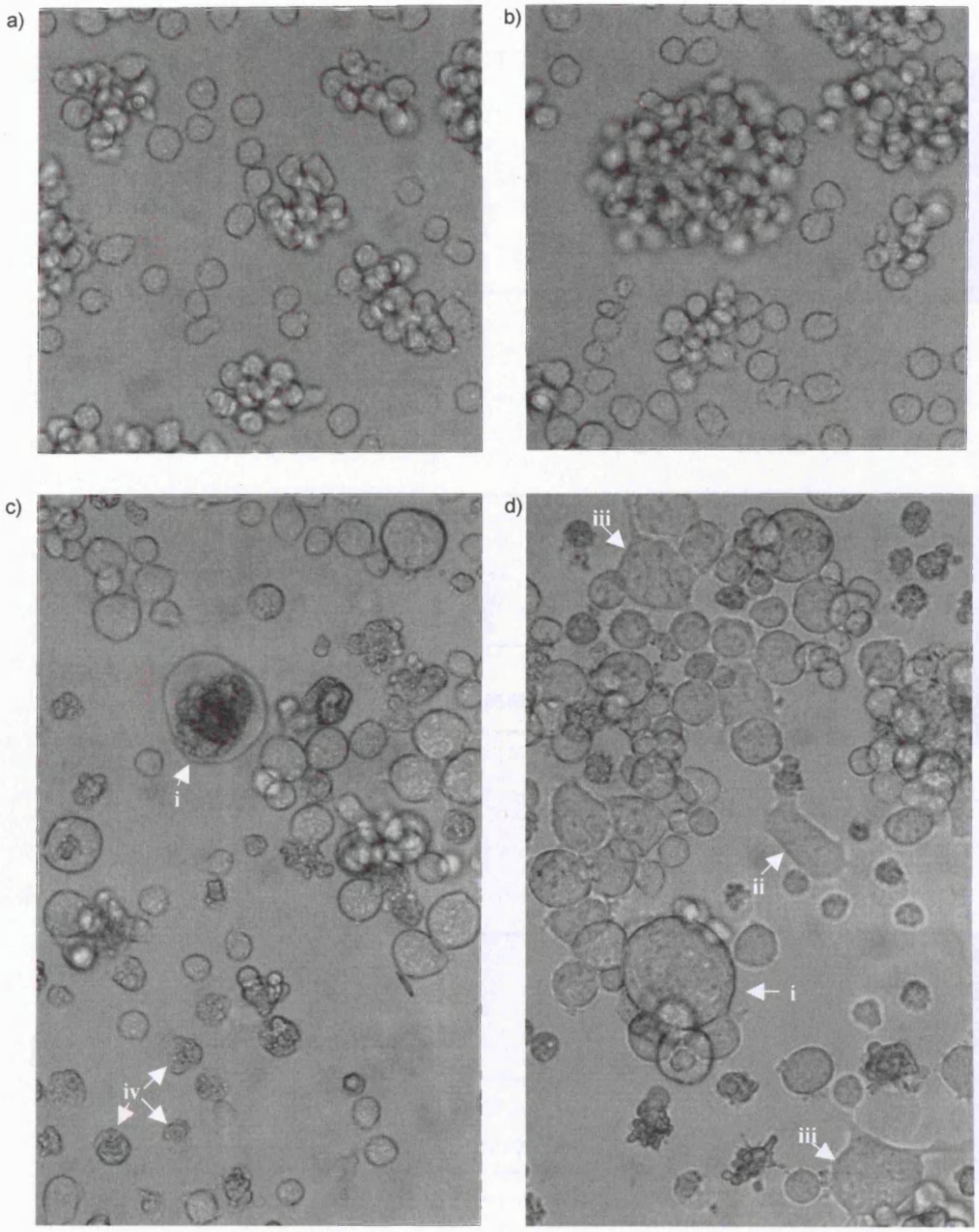


Figure 4.1 Confocal microscopy photograph of unstained Sup-T1 cells. a) uninfected cells; b) cells treated with uninfected cell lysate; ; c-d) HHV-7 infected cells showing cytopathic effect: giant multinucleated syncytia (i); flattened, elongated cells loosely attached to bottom of cell culture flask (ii); cells that are both multinucleated and attached (iii), and cells with plasma membrane 'bubbles' (iv).

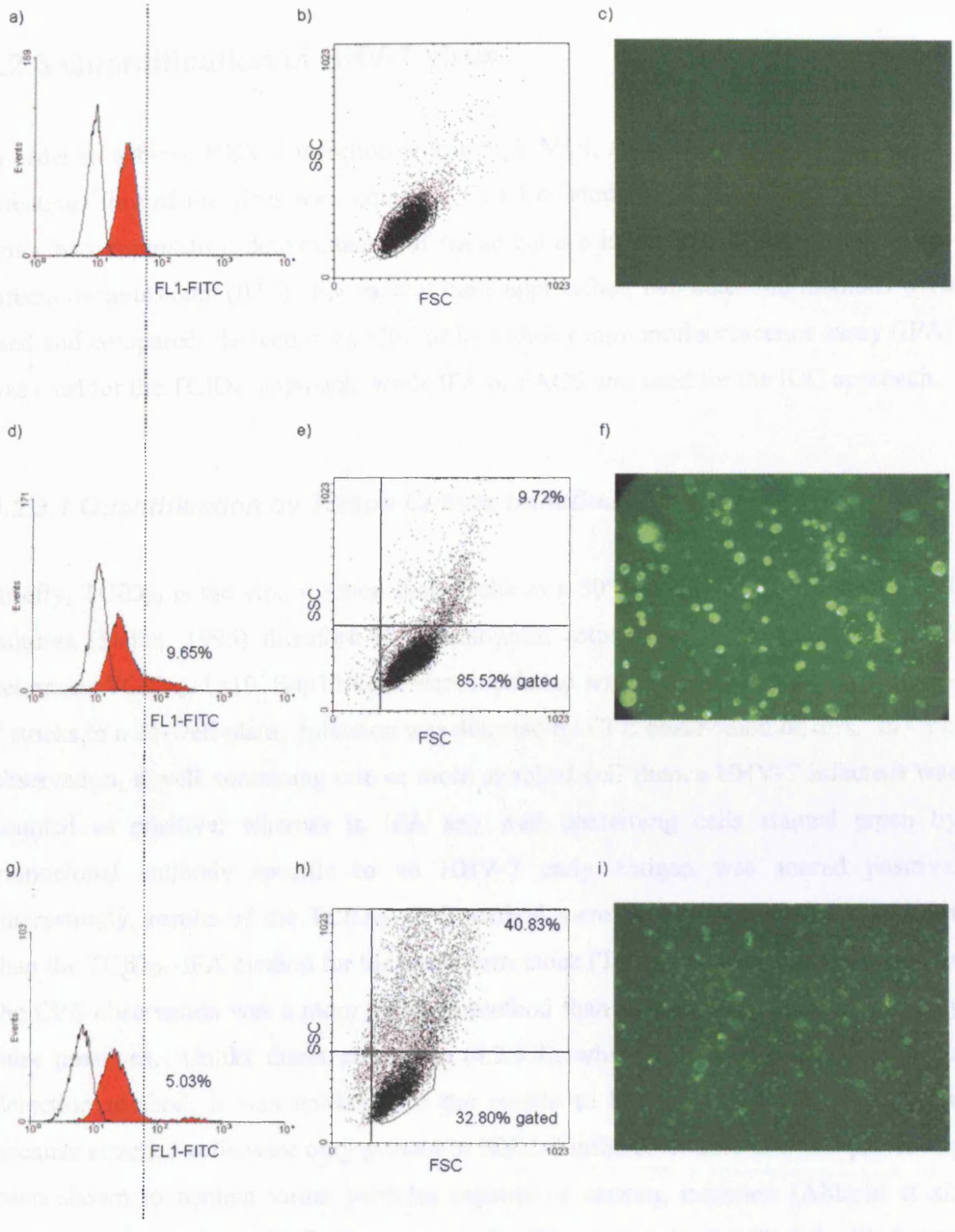


Figure 4.2 Flow cytometry and immunofluorescence of uninfected SupT1 (a-c), SupT1 infected with cell-free HHV-7 virus (d-f), and HHV-7 infected SupT1 cells maintained by co-culture (g-i). Percentage of large and granular cells (upper right quadrant) in dot plots (e,h) do not correlate with percentage of HHV-7 antigen-bearing cells (FL1-FITC greater than 60, dotted line in panels d,g). FSC, forward scatter; SSC, side scatter; red histogram, stained with anti-HHV-7 antibody; white histogram, unstained

4.2.3 Quantification of HHV-7 virus

In order to achieve HHV-7 infection with a high MOI, an accurate way of measuring infectious titre of the virus was required. Cell free stocks of HHV-7 were quantified using two approaches: determination of tissue culture infectious $dose_{50}$ ($TCID_{50}$) and infectious units count (IUC). For each of these approaches, two detection methods were used and compared. Detection by CPE or by indirect immunofluorescence assay (IFA) was used for the $TCID_{50}$ approach; while IFA or FACS was used for the IUC approach.

4.2.3.1 Quantification by Tissue Culture Infectious $Dose_{50}$ ($TCID_{50}$)

Briefly, $TCID_{50}$ is the viral dilution that results in a 50% chance of infection in target cultures (Peden, 1995) therefore is an end-point estimation of the infectivity. To determine $TCID_{50}$, 1×10^5 SupT1 cells were incubated with serial dilutions of two HHV-7 stocks in a 24-well plate. Infection was detected by CPE observation or IFA. In CPE observation, a well containing one or more attached cell (hence HHV-7 infected) was counted as positive; whereas in IFA any well containing cells stained green by monoclonal antibody specific to an HHV-7 early antigen was scored positive. Interestingly, results of the $TCID_{50}$–CPE method were approximately one log higher than the $TCID_{50}$–IFA method for the same virus stock (Table 4.1), suggesting that either the CPE observation was a more sensitive method than IFA, or that it was affected by false positives. Unlike direct estimation (4.2.3.2), when CPE was used as end-point detection method, it was unlikely for the results to be affected by false positives, because attached cells were only present in HHV-7 infected culture and had previously been shown to contain virion particles capable of causing infection (Ablashi et al., 1998). Also, no attached cells were seen in SupT1 cultures mock-infected with freeze-thawed lysate of uninfected cells (Figure 4.1b). Therefore we believe that the $TCID_{50}$–CPE method was a robust measure of virus infectious titre. Another possible reason for the different results might be because the CPE observation was made at 6 dpi, but cells were harvested for IFA at 4 dpi. Sensitivity of the $TCID_{50}$–IFA method may be improved by growing the cells for a longer period before staining to allow extra time for

antigen production. However, it was unknown whether this would account for the difference between $TCID_{50}$–CPE and $TCID_{50}$–IFA results.

4.2.3.2 Quantification by Infectious units count (IUC)

In the second approach, IUC was calculated from the percentage of infected cells bearing HHV-7 antigen, which was detected either by IFA (Figure 4.3) or FACS (Figure 4.4) using the anti-HHV-7 antibody and the results translated into infectious units per ml (Table 4.1). The anti-HHV-7 antibody specifically bound to HHV-7 infected cells, but not to uninfected cells or cells infected by HHV-6, a close relative of HHV-7 (Figure 4.3). Therefore this IUC–IFA method had a low false positive rate and was considered a highly specific method for detecting HHV-7 infected cells. Results obtained were approximately 3 times higher than $TCID_{50}$–CPE results (Table 4.1 & Figure 4.6).

IUC–FACS was the only method to show that stock 020723 had a higher titre than stock 040520 and the difference was about 20 times (Table 4.1). Note that only the major population of small cells (within gated region) were taken into account (Figure 4.4), whereas any HHV-7 infected cells in the large and granular population (outside gate) were excluded. The gating was necessary to remove false positives, because large and/or granular cells in HHV-7 infected cultures displayed high background fluorescence even in the absence of anti-HHV-7 antibodies (Figure 4.5). The drawback for this filtering was that the proportion of fluorescent cells in the gated region may be different from that in the whole cell population, which may account for the low estimate for stock 040520. Perhaps most of the cells infected by stock 040520 became larger and more granular, hence excluded by FACS gating, whereas the cells infected by stock 020723 remained small. Unfortunately, it was not possible to confirm this by FACS, due to the high background fluorescence of the large, granular cells (Figure 4.5). In contrast, the human eye was more accurate in differentiating the antigen-bearing cells, by their characteristic punctate staining pattern, from the background fluorescence (Figure 4.3). Of the two methods for infectious unit count, IUC–IFA was therefore considered more accurate than IUC–FACS.

Table 4.1 Results using different quantification methods of HHV-7 stocks

HHV-7 stock	$TCID_{50}$ / ml		Infectious units / ml	
	by CPE[a]	by IFA[b]	by FACS[c]	by IFA[c]
020723	3.44E+04	1.81E+03	4.36E+05	8.44E+04
040520	4.87E+04	7.24E+03	1.90E+04	1.33E+05

[a] $TCID_{50}$ value determined by scoring positive wells by the presence of attached cells

[b] $TCID_{50}$ value determined by scoring positive wells by the presence of fluorescent cells

[c] percentage of HHV-7 antigen-bearing cells × number of cells/ volume of virus

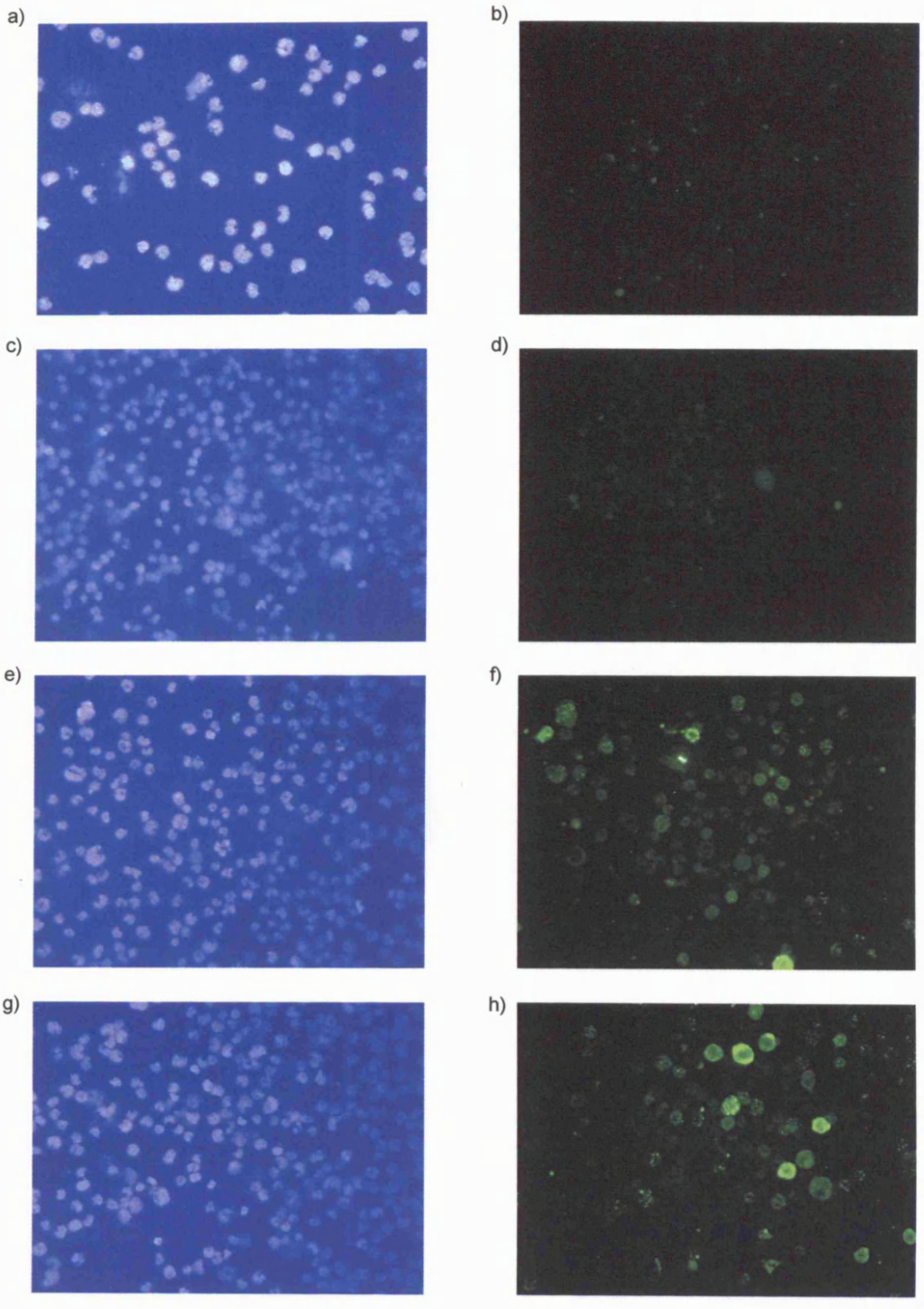


Figure 4.3 IFA for HHV-7 infection. SupT1 cells uninfected (a-b); HHV-6 infected (c-d); HHV-7 stock 020723 infected (e); HHV-7 stock 040520 infected (f) were fixed and stained with anti-HHV-7 monoclonal antibody, followed by FITC to stain HHV-7 antigen-bearing cells (b,d,f,h); and DAPI counter-stain to visualise all cells (a,c,e,g)

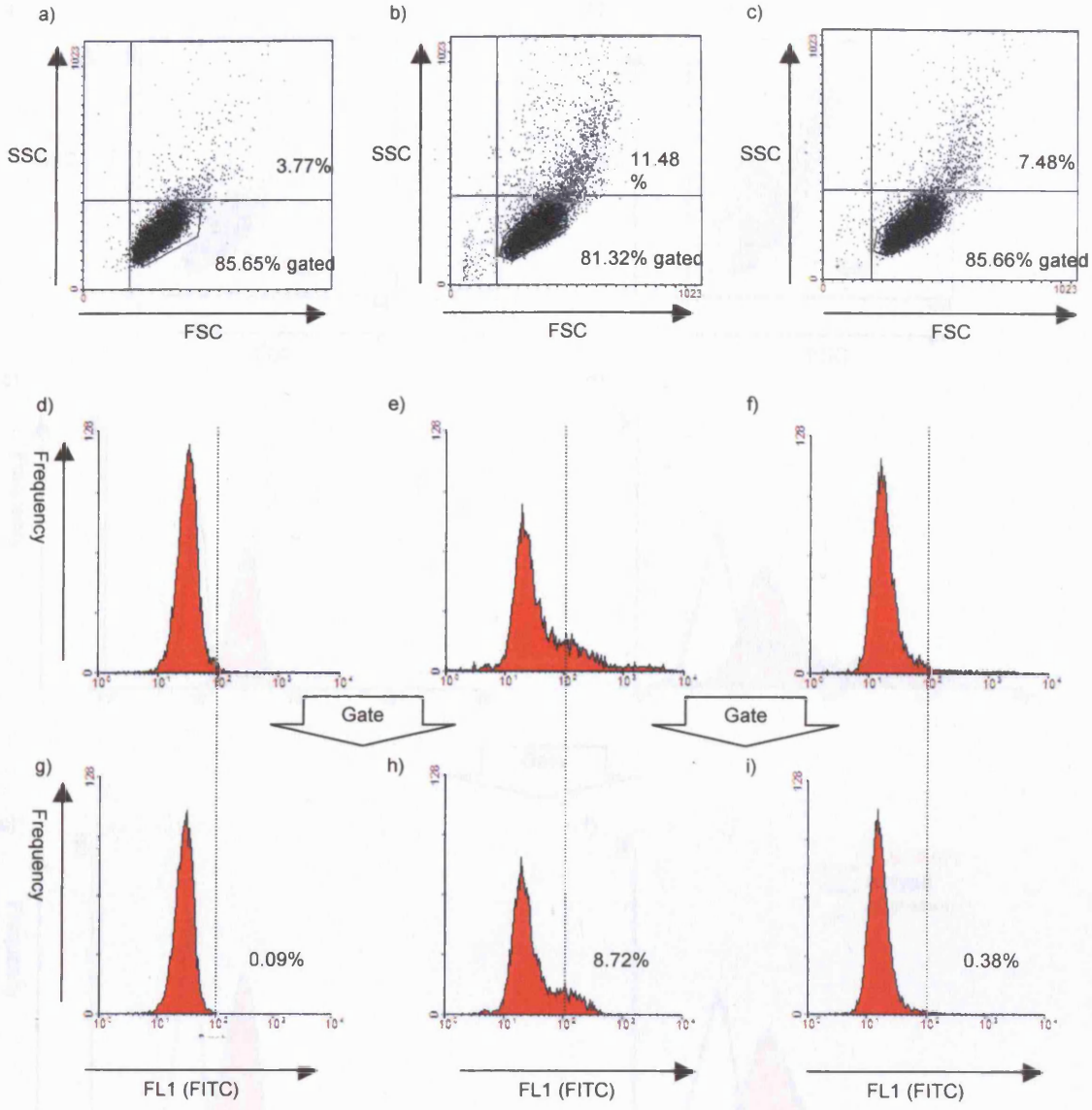


Figure 4.4 Flow cytometry analysis of SupT1 uninfected (a,d,g); infected with HHV-7 stock 020723 (b,e,h); or stock 040520 (c,f,i). Results are representative of triplicate measurements. Infection was estimated using the percentage of HHV-7 antigen-bearing cells (FL1-FITC greater than 10^2, dotted line) in the gated region (g-i). See also the legend for Figure 4.5

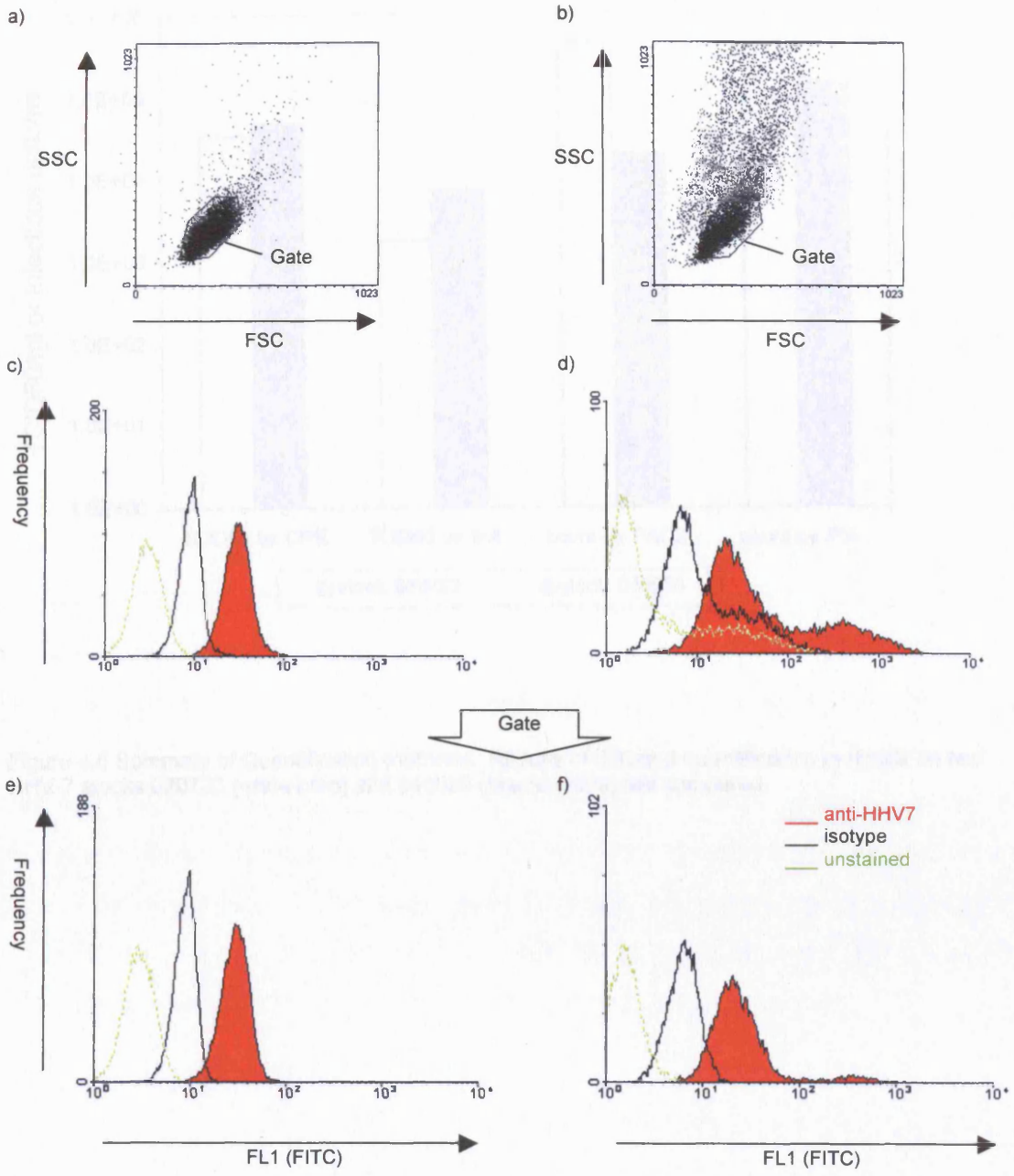


Figure 4.5 Flow cytometry analysis of uninfected (a,c,e) and HHV-7 infected SuptT1 cells (b,d,f). Infected cells (b) contained more larger (FSC) and granular (SSC) cells than uninfected cells (a). Without gating, infected cells displayed a 'tail' portion (d) of higher fluorescence (FL1 FITC) when stained with anti-HHV-7 antibody (red shaded), secondary isotype control (black curve), or unstained (green curve). Gating the population of smaller cells (black polygon) completely removed these tails in unstained and isotype controls, and reduced the tail in anti-HHV-7 stained cells (f). Gating has no effect on the fluorescence of uninfected cells (c,e)

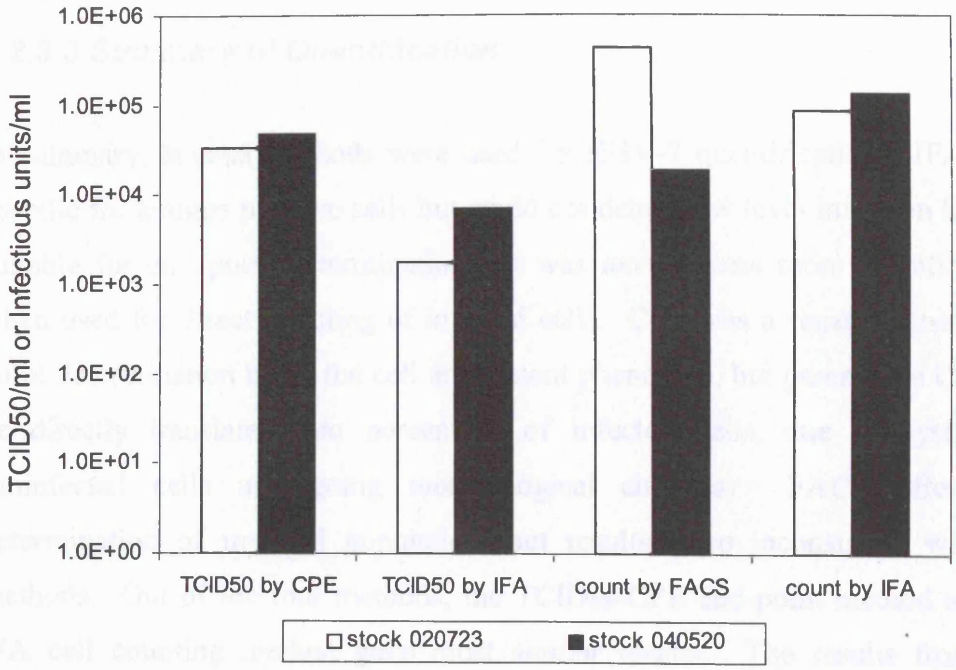


Figure 4.6 Summary of Quantification methods. Results of different quantification methods on two HHV-7 stocks 020723 (white bars) and 040520 (shaded bars) are compared.

4.2.3.3 Summary of Quantification

In summary, several methods were used for HHV-7 quantification. IFA was highly specific for antigen positive cells but could not detect low level infection hence was not suitable for end point determination. It was nevertheless more specific than FACS when used for direct counting of infected cells. CPE was a sensitive method for end-point determination using the cell attachment phenotype, but percentage CPE could not be directly translated into percentage of infected cells, due to bystander effect. (uninfected cells undergoing morphological changes) FACS offered a quick determination of infected population, but results were inconsistent with the other methods. Out of the four methods, the $TCID_{50}$–CPE end-point method and the IUC–IFA cell counting method gave most similar results. The results from these two methods had a reproducible difference of 2.5-fold (Figure 4.6 and Table 4.1) in both virus stocks tested. This systematic difference was probably because one method calculates $TCID_{50}$ value and the other measures infectious units. This also suggested that both methods produced reproducible results, therefore were believed to be the most accurate methods. Subsequent virus quantification experiments were therefore based on these two approaches. This thesis did not apply other methods of assessing viral infection, such as the detection of viral DNA by southern blotting, PCR or real-time PCR; viral proteins by ELISA; or viral particles by electron microscopy.

4.2.4 Optimisation of HHV-7 culturing to achieve a high multiplicity of infection

In order to achieve high multiplicity of infection for the gene expression studies, it was desirable to produce virus stocks of high infectious titre and to optimise the efficiency of infection. Having established accurate methods for HHV-7 titration, we were able to examine the effect of cell culture conditions on the titre of HHV-7 virus stocks.

4.2.4.1 Production of high tire virus

The first attempt to grow high titre HHV-7 virus was to expand the virus culture (in SupT1 cells) into multiple large volume (100 ml) flasks as described in section 2.3.4. When 15-30 flasks (1500 – 3000 ml) of HHV-7 infected cells (5 – 25% CPE) were harvested, the virus was released by freeze-thawing in $1/100^{th}$ volume as described in section 2.3.4. This meant 7.5×10^7 - 7.5×10^8 CPE-displaying cells were concentrated to a volume of 15 – 30 ml, resulting in 5.0×10^6 - 2.5×10^7 CPE-displaying cells per ml lysate (Table 4.2). This method has been used previously (Menegazzi et al., 1999), but the titres of virus stocks we obtained using this method were rather low (average titre = 1.9×10^3 infectious units/ml) (Table 4.2). In light of this, HHV-7–infected cultures were harvested in medium (300 ml) and small (6 – 10 ml) volumes, and their titres determined and compared (Table 4.2). Results showed that cells (40-75% CPE) grown in small volumes (6-9 ml) for freeze-thawing in 1-2 ml gave an average titre of 1.2×10^5 infectious units/ml (Table 4.2). This was almost 100 times higher than that obtained using the large volume method.

Table 4.2 Production of high titre virus stocks

HHV-7 Culture volume (ml)	Cells/ml	% CPE	No. CPE-displaying cells [a]	Lysate volume (ml)	CPE/ml lysate [b]	Mean titre (infectious units/ml) [c]
1500-3000	1×10^6	5-25%	7.5×10^7 - 7.5×10^8	15-30	5.0×10^6 - 2.5×10^7	1.9×10^3
300	1×10^6	25-50%	7.5×10^7 - 1.5×10^8	30	2.5×10^6 - 5.0×10^6	2.0×10^4
6-10	1×10^6	40-75%	2.4×10^6 - 7.5×10^6	1-2	2.4×10^6 - 3.8×10^6	1.2×10^5

[a] No. of CPE-displaying cells = culture volume × no. cells/ml × % CPE

[b] CPE/ml lysate = no. of CPE-displaying cells / lysate volume

[c] For large volume (1500-3000ml), mean titre of 4 different stocks; for medium volume (300ml), mean titre of 2 different stocks; for small volume (6-10ml), mean titre of 7 different stocks; all calculated using the IUC-IFA quantification method

The reason for this was probably because it was more efficient for a culture in a small volume to reach high CPE percentage (40-75% at 6-8 days post co-culture); whereas cultures in large volumes often had lower CPE (5-25%), perhaps because a larger culture volume made cell-to-cell spread of viral infection more difficult. However, if the CPE estimation was correct, each ml of virus stock from a 1500 ml culture at 5% CPE would still contain more CPE-displaying cells than that from a 10 ml culture at 75% CPE (Table 4.2). This finding suggested that CPE of the culture was not directly related to number of infectious particles produced. HHV-7 was subsequently cultured in multiple (10-20) flasks of small volumes (6-10 ml) and harvested as 1-2ml stocks when the cells reached a CPE of 40-75%, as we have found that this provided stocks with the highest titre. The stocks were then pooled and quantified (section 2.3.5) before use.

There could be another possible reason for the unproportionally low degree of CPE in the large volume cultures. As the virus was passaged in SupT1 cells by co-culturing, the virus might have established latent infection in some cells, resulting in a loss of CPE and reduced production of infectious virus as passage number increased. Subsequent lysis of the cells and virus stock production therefore yielded less infectious virus per ml. In contrast, in the small volume culturing method, viruses were harvested as soon as the small culture reached an adequate CPE, thus avoiding the loss of productive infection with passaging.

4.2.4.2 Effect of Spinoculation & Polybrene on Inoculation efficiency

After optimisation of virus stock preparation, infectious titres of approximately 1×10^5 infectious units/ml were achieved (section 4.2.4.1). The next step was to improve the efficiency of HHV-7 virus inoculation, thereby increasing the effective MOI. Firstly, 1×10^5 SupT1 cells were inoculated with 100 µl of the virus stock 020723 for 1 hour at 37 °C. After the inoculation period, the virus supernatant was removed and used to inoculate fresh 1×10^5 SupT1 cells, in order to determine the amount of any unadsorbed virus remaining and if these were able to cause further infection. Fresh medium was added to the originally inoculated cells, and all cells were cultured at 37°C for 3 days, after which they were harvested, washed, and fixed onto slides for infectious unit count by IFA.

Results (Table 4.3) indicated that 1 ml of stock 020723 virus (estimated infectious titre $3\text{-}8\times10^4$ infectious units/ml, Table 4.1) was able to infect 6.8×10^4 cells after a 1-hour incubation period at 37°C, and the remaining virus supernatant was able to infect a further 1.4×10^5 cells, hence a total of 2.1×10^5 cells. This suggested the virus only achieved 33% of total possible infection after 1h incubation. This value coincided with the titre obtained for stock 020723, which was also determined after inoculation. However, this also showed that much infectious virus remained after 1 hour inoculation. Therefore methods to increase infection efficiency were investigated. Infection by spinoculation at 1000 g, 20 °C for 1 hour increased the number of infected cells by about 3.5-fold to 2.3×10^5 cells. Interestingly, the remaining virus only infected 1.6×10^4 more cells, bringing the total number of infected cells to 2.5×10^5 cells/ ml virus, similar to the total infection obtained without spinoculation. Therefore spinoculation for 1 hour achieved about 94% of total possible infection. If 2.5×10^5 cells/ ml virus was the maximum achievable infection of stock 020723, then the inoculation efficiency of 1-hour incubation and 1-hour spinoculation was 27% and 94%, respectively (Table 4.3 & Figure 4.7a). These findings suggested that spinoculation was a more efficient inoculation method for HHV-7 infection of SupT1 cells.

The effect of polybrene on HHV-7 infection of SupT1 cells was also investigated. Results showed that the presence of polybrene (8 µg/ml) during spinoculation led to a

decreased number of infected cells to 4.3×10^4 cells/ml virus, $1/5^{th}$ of the number infected when spinoculation was used alone (Table 4.3 & Figure 4.7a). When polybrene was used during incubation at 37 °C without spinoculation, no infected cells were detected at all. Therefore the presence of polybrene (8 μg/ml) had an inhibitory effect on HHV-7 infection of SupT1 cells with or without spinoculation.

Table 4.3 Effects of spinoculation and polybrene on HHV-7 stock 020723

	Number of infected cells / ml virus				
	Original inoculum	Remaining supernatant	Total[a]	% Total[b]	% Maximum[c]
Incubation	6.8×10^4	1.4×10^5	2.1×10^5	33	27
Spinoculation	2.3×10^5	1.6×10^4	2.5×10^5	94	94
Incubation + polybrene	0	0	0	n/a	n/a
Spinoculation + polybrene	4.3×10^4	1.4×10^4	5.7×10^4	75	17

[a] Total = infected cells/ml virus in original inoculum + infected cells/ml remaining virus in supernatant
[b] % Total = infected cells/ml virus in original inoculum / total infected cells/ml virus achieved by the particular inoculation method × 100
[c] % Maximum = infected cells/ml virus in original inoculum / total infected cells/ ml virus achieved by spinoculation × 100

In order to consolidate these findings, the experiment was repeated with several other virus stocks (Table 4.4). Again, spinoculation resulted in higher numbers of infected cells compared to incubation, with fold increases of 2.1, 2.5, 1.1, 3.4, and 3.8 for stocks 050118A, B, D, F and G, respectively, (Figure 4.7b). Except for stock 050118G, spinoculation also resulted in a higher percentage (91-100%) of total infection, compared to incubation alone (30-58%), as observed before in virus stock 020723 (Table 4.4). Moreover, for stocks 050118A-D, there were no detectable viruses in the remaining supernatant after spinoculation. Although the different stocks had varying infectious titres, spinoculation at 1000 g increased the infectious titres of all stocks (except stock 050118D) by 2.1- to 3.8-fold, clearly demonstrating that spinoculation was the more effective method. The enhancement effect of spinoculation on HHV-7 infectivity was consistent with previous studies in HIV (Ho et al., 1993; O'Doherty et al.,

2000; Pietroboni et al., 1989), HSV (Tenser and Dunstan, 1980), HHV-8 (Inoue et al., 2003), CMV (Ho et al., 1993; Hudson, 1988; Hudson et al., 1976) and HHV-6 (Pietroboni et al., 1989). However, our results showed that HHV-7 infectivity was only enhanced slightly, compared to 10-fold in HIV and 100-fold in HHV-6 (Pietroboni et al., 1989).

On the other hand, the presence of polybrene (8 μg/ml) had a negative effect on HHV-7 infectivity with or without spinoculation (Table 4.3, 4.4), consistent with stock 020723 and published data on HHV-6 (Black et al., 1989), but not with previous studies on HHV-8 (Inoue et al., 2003) and HIV (Pan et al., 1993). Perhaps the presence of polybrene had adverse effects on SupT1 cells proliferation, therefore inhibiting secondary infection by cell-to-cell spread.

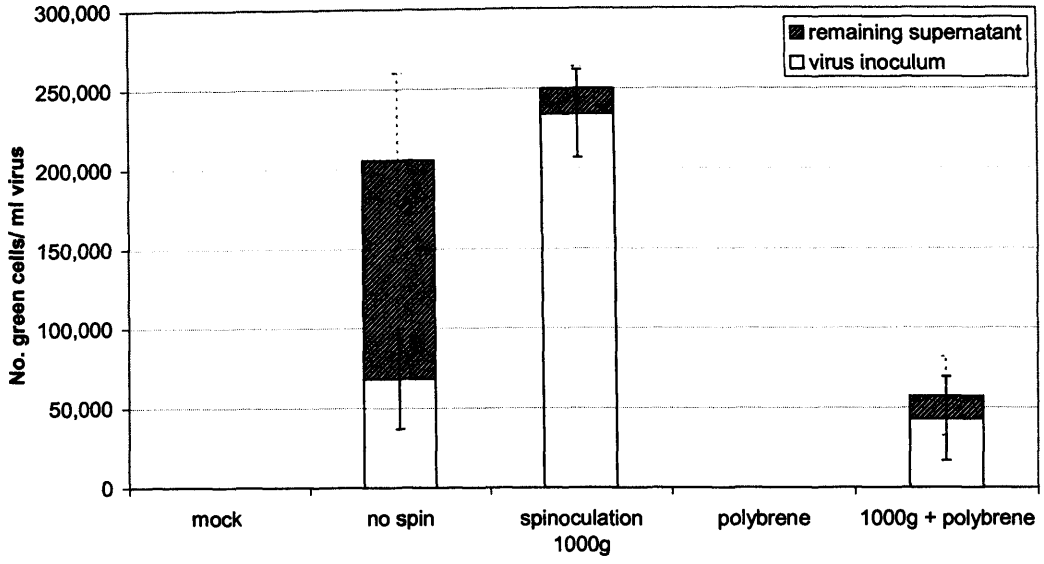
Therefore spinoculation in the absence of polybrene will be the method of choice for inoculating SupT1 cells with HHV-7 virus in subsequent experiments. From these data we have shown that, with the HHV-7 virus stock production method available (section 4.2.4.1), we can therefore achieve an MOI of 2.5 on 1×10^5 SupT1 cells, giving a theoretically synchronous infection.

Table 4.4 Effects of spinoculation and polybrene on various HHV-7 virus stocks

	Number of infected cells / ml virus and % of total possible infection							
	No Spin		Spinoculate 1000g		Polybrene		1000g + Polybrene	
Stock	Original inoculum	% Total[a]	Original inoculum	% Total[a]	Original inoculum	% Total[a]	Original inoculum	% Total[a]
020723	6.8×10^4	33	2.3×10^5	94	0	n/a	4.3×10^4	75
050118A	1.5×10^4	30	3.0×10^5	100	0	n/a	0	n/a
050118B	1.6×10^5	43	4.0×10^5	100	0	n/a	0	n/a
050118D	1.5×10^5	58	1.7×10^5	100	0	n/a	0	n/a
050118F	9.3×10^4	37	3.2×10^5	91	0	n/a	0	n/a
050118G	6.7×10^4	70	2.5×10^5	55	0	n/a	1.6×10^5	100

[a] % Total = infected cells/ml virus in original inoculum / total infected cells/ml virus achieved by the particular inoculation method × 100. 100% means there were no detectable infectious virus left in the supernatant. For simplicity the number of cells infected by the remaining virus supernatant is not displayed. Also see notes for Table 4.3

a)



b)

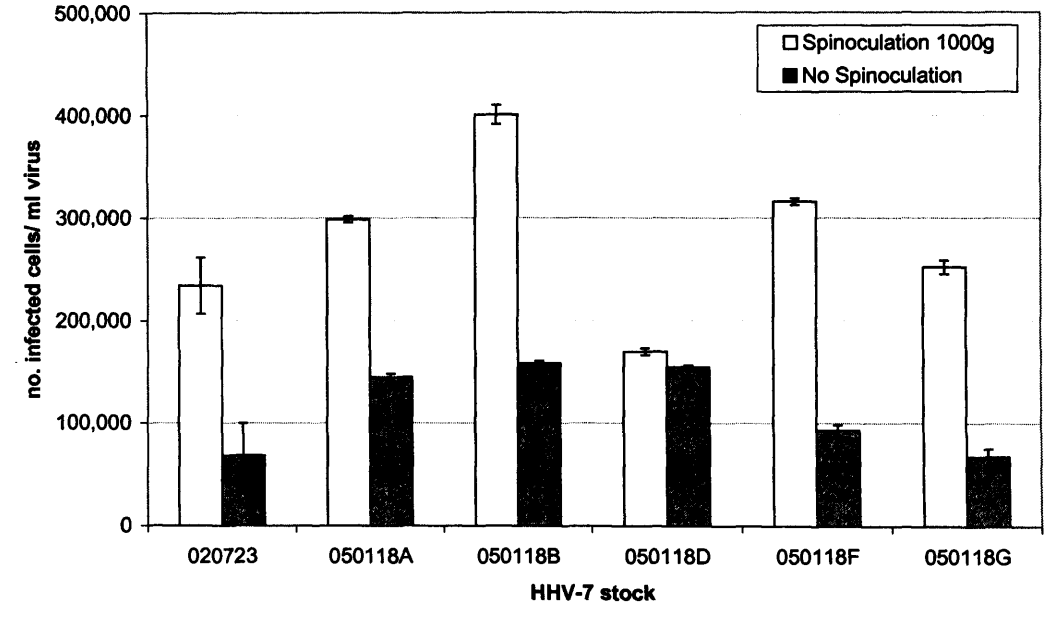


Figure 4.7 Effect of spinoculation on HHV-7 infection. a) 10^5 Sup-T_1 cells were inoculated with 0.1 ml of virus stock 020723 for 1 hour either at 37°C (no spin), with spinoculation (1000g, 20°C), with polybrene (8 ug/ml), or with both spinoculation and polybrene. After inoculation period, leftover virus supernatant was removed and added to fresh Sup-T_1 cells (striped bars). Fresh medium was added to the originally inoculated cells (white bars). All cells were harvested after 3 days, and the number of infected cells determined by IFA with an anti-HHV-7 antibody. Results represent the mean of multiple measurements and the error bars represent the standard deviation. b) Experiment was repeated for 5 different HHV-7 stocks, only data for spinoculated cells (white bars) and incubated cells without spinoculation (shaded bars) are shown.

4.2.5 RNA amplification

4.2.5.1 Reasons for RNA amplification

Despite attempts (described in sections 4.2.3 and 4.2.4) to improve the infectious titre of HHV-7 virus stocks, the highest titres achieved were only $2\text{-}4\times10^5$ infectious units/ml virus (with spinoculation), whereas at least 5×10^6 cells were needed to give enough RNA for each array hybridisation. When spinoculation was used on 5×10^6 cells (MOI of ~0.05), no CPE was visible at 48 hpi. When gene expression of these cells were analysed, the mean fluorescence signal and pass-rate for the HHV-7 probes in the Cy5 channel were low (Table 4.5), and the expression pattern probably represented the majority of uninfected cells rather than the infected cells. Given the limited virus titre, the only available option was to decrease the number of cells used in infection to 1×10^5 cells, so that a higher MOI of 2.5 (section 4.2.4.2) could be achieved. Since the amount of total RNA extracted from such a low cell number (0.5 – 2.5 μg) were below the amount normally required for DNA microarray analysis (50 – 200 μg), amplification of the RNA was required. Conventional methods produced 3'-biased, antisense amplified RNA (aRNA), ranging in size between 300-3000 nucleotides, with a peak at 600 nucleotides (Figure 4.9b). These 3'-biased aRNAs were too short to encompass some of the HHV-7 array probes, which were designed to target the 5'-end of HHV-7 ORFs (Figure 3.2). Therefore sense strand RNA amplification was needed to produce viral aRNA from the 5'-end.

4.2.5.2 Sense-strand amplification

Total RNA from 1×10^5 SupT1 cells infected by HHV-7 was amplified by the sense strand RNA amplification method (Figure 4.8, also see section 2.7.2.2). Briefly, first strand cDNA was synthesised from the 3'-end of mRNA by reverse transcription using a mixture of oligo-dT and random primers, in a reaction to favour full length cDNA synthesis. A poly-dT tail was added to the 3'-end of the cDNA, which was used for the annealing of a T7 primer. The 3'-end of the cDNA was then 'filled in' by Klenow fragment to generate a double stranded T7 promoter sequence. From this T7 promoter

site, sense strand amplified RNA copies were produced by *in vitro* transcription using T7 RNA polymerase. This method produced amplified sense strand RNA (aRNA) of high yield, with an average of 5.6 μg aRNA from 250 ng total RNA. Preliminary results from 12 mRNA purifications (using a column to bind polyA-tailed mRNA, section 2.7.1.3) suggested that total RNA in SupT1 cells was made up of, on average, 1.8% of mRNA, therefore sense strand RNA amplification using a combination of oligo-dT and random primers achieved on average a 1250-fold amplification of mRNA into aRNA. The yield was much higher than that attained when oligo-dT primer was used alone (i.e. no random primers), which only produced on average 2.8 μg aRNA from 2 μg total RNA (77-fold amplification). The aRNA products ranged from 100 to 4000 nucleotides in length, peaking at 200-500 nucleotides (Figure 4.9c), which was sufficient for recognition by the HHV-7 and human gene probes located at the 5'-end of transcripts.

Hybridisation of the fluorescent-labelled aRNAs (converted to cDNAs by reverse transcription) to the custom-made HHV-7 microarrays (see Chapter 3) produced satisfactory results. For example, an array (S19) was hybridised to Cy5-labelled aRNA (500 ng) from HHV-7 infected cells (48 hpi) and Cy3-labelled reference aRNA. The cDNAs bound to the array probes and gave good fluorescent signals, with 71.9% and 68.6% of the probes (viral and human) passing the signal-to-noise filtering criteria (see section 3.2.4.2) in the Cy5 and Cy3 channels, respectively (Table 4.5). All the HHV-7 probes had signal-to-noise ratios (SNR) above the filtering threshold, a pass rate of 100% in both channels, while the blank spots containing only buffer had low SNRs and had a pass rate of 0%. The Cy5 mean signal of the HHV-7 array probes was more than 5 times higher than the all-probes mean. It was unlikely for this difference to be due to a bias introduced by RNA amplification (Schindler et al., 2005; Wilson et al., 2004), because such difference was not observed in reference RNA, which was also amplified. Rather, the higher HHV-7 signals may reflect a large population of HHV-7 aRNAs in the Cy5-labelled 48 hpi sample. Moreover, the signal and pass-rate of HHV-7 probes were much (approximately 30 times) higher in the aRNA array than in the unamplified RNA array (Table 4.5), indicating that the aRNAs contain more copies of the HHV-7 transcripts. The aRNA and unamplified RNA were extracted from 1×10^5 and 5×10^6 infected SupT1 cells, respectively. Therefore the reduction of cells used in infection,

followed by sense strand RNA amplification (using a combination of oligo-dT and random primers), had improved the detection of HHV-7 transcripts.

Table 4.5 Example arrays hybridised to messenger RNA or amplified RNA

Sample	Spot group	Cy5		Cy3		Cy5 SNR pass rate (%)	Cy3 SNR pass rate (%)
		Mean signal intensity[a]	Mean SNR[b]	Mean signal intensity[a]	Mean SNR[b]		
Messenger RNA[d]	HHV-7 blanks	74.0	1.21	388.6	1.46	0.0	0.0
	HHV-7 ORFs	115.5	1.90	816.4	3.13	10.2	100.0
	All probes[c]	383.9	4.85	1293.8	4.96	76.6	70.5
Amplified RNA[e]	HHV-7 blanks	56.9	0.93	194.3	0.73	0.0	0.0
	HHV-7 ORFs	3645.3	57.07	903.6	3.49	100.0	100.0
	All probes[c]	644.3	10.58	873.8	3.35	71.9	68.6

[a] Mean signal intensity = mean of the background-subtracted signal intensities of all the spots in the particular spot group, in one channel

[b] SNR = signal-to-noise ratio = (background-corrected signal)/background

[c] All spots containing array probes were included; but empty spots were excluded

[d] Messenger RNA from 5×10^6 HHV-7 infected cells (48hpi) and Reference RNA were converted into Cy5- and Cy3-labelled cDNA, respectively

[e] Messenger RNA from 1×10^6 HHV-7 infected cells (48hpi) and Reference RNA were sense strand-amplified and then converted into Cy5- and Cy3-labelled cDNA, respectively

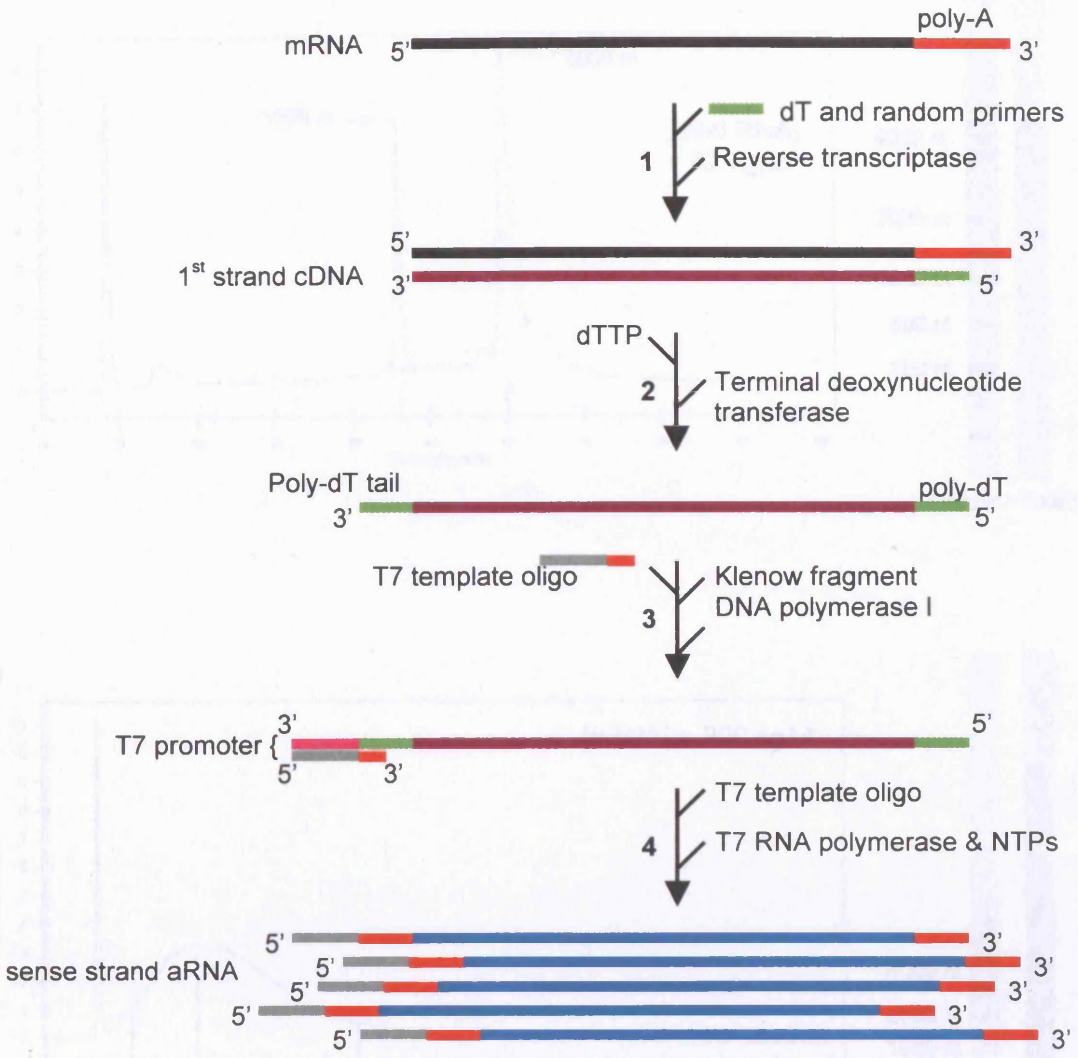


Figure 4.8 . Sense-strand amplification of RNA (adapted from SenseAmp, Genisphere) Messenger RNA were amplified by: 1) first strand cDNA synthesis by reverse transcription using oligo-dT and random primers; 2) poly-dT tailing of cDNA by terminal deoxynucleotide transferase; 3) annealing of a T7 template oligo, followed by T7 promoter synthesis using the 'filling in' property of Klenow fragment; 4) *In vitro* transcription of sense strand aRNA from the T7 promoter by T7 RNA polymerase.

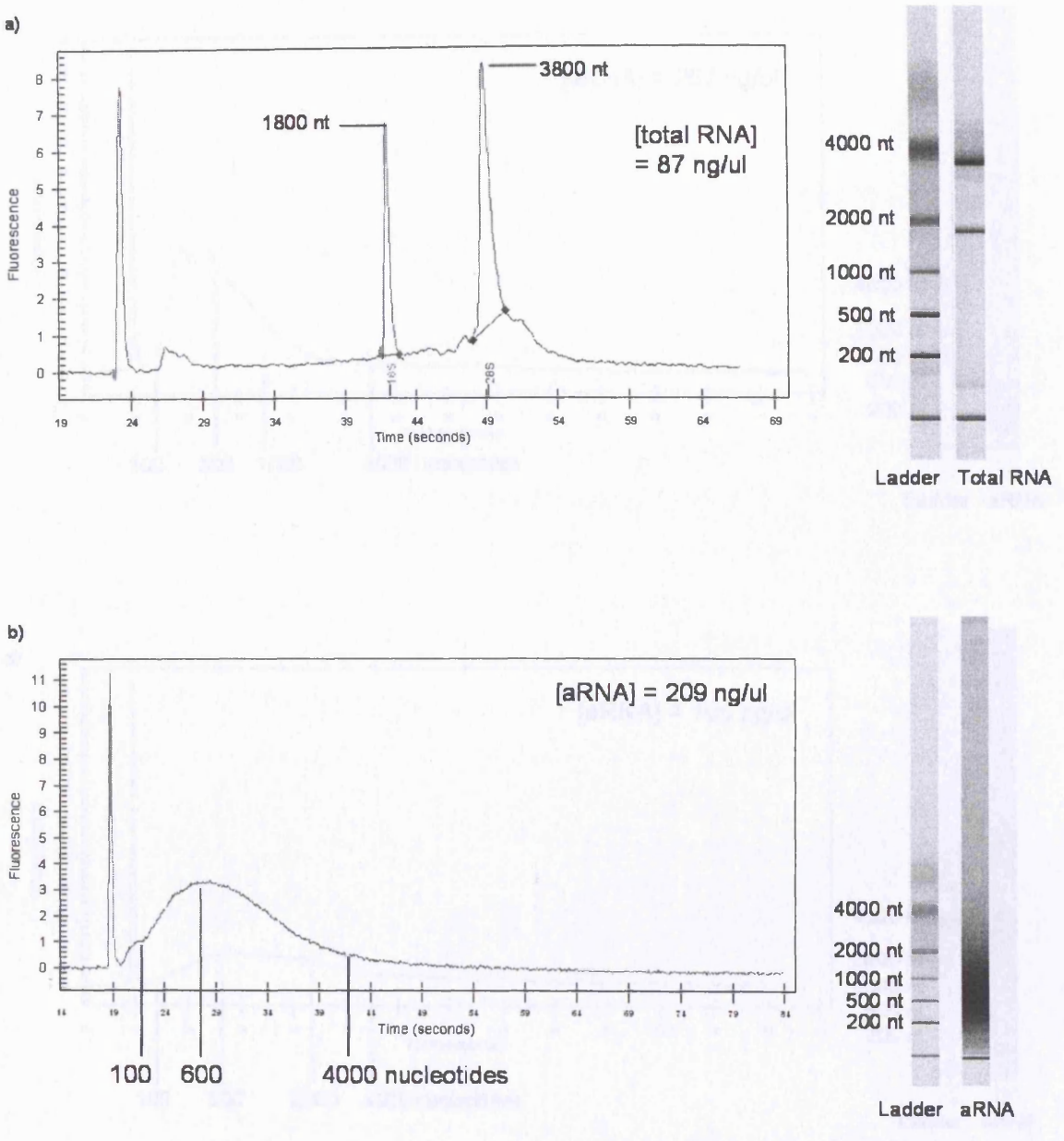


Figure 4.9 From 1 $\times 10^5$ SupT1 cells, total RNA and amplified RNA (aRNA) were analysed by Agilent Bioanalyser. a) total RNA; b) antisense aRNA; next page: c) sense strand aRNA amplified using oligo-dT and random primers; and d) sense strand aRNA amplified using oligo-dT primer only. The two ribosomal RNA species (18S and 28S) can be seen in total RNA (both graph and gel diagram), but absent in aRNA. [total RNA] and [aRNA] denotes the concentration. The sizes of aRNA products ranged between 100-4000 nucleotides.

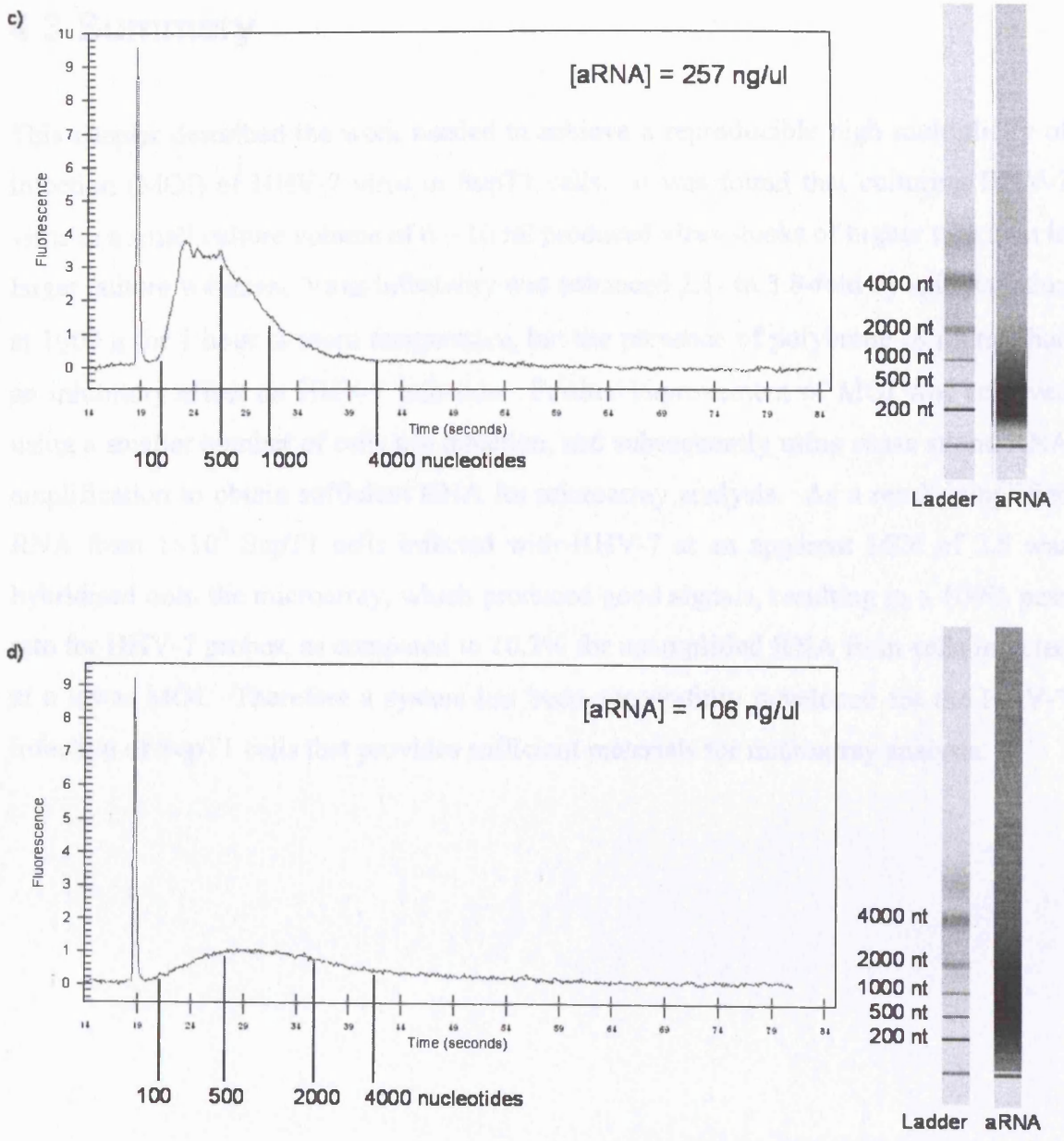


Figure 4.9 (continued)

4.3 Summary

This chapter described the work needed to achieve a reproducible high multiplicity of infection (MOI) of HHV-7 virus in SupT1 cells. It was found that culturing HHV-7 virus in a small culture volume of 6 – 10 ml produced virus stocks of higher titre than in larger culture volumes. Virus infectivity was enhanced 2.1- to 3.8-fold by spinoculation at 1000 g for 1 hour at room temperature, but the presence of polybrene (8 μg/ml) had an inhibitory effect on HHV-7 infection. Further improvement of MOI was achieved using a smaller number of cells per infection, and subsequently using sense strand RNA amplification to obtain sufficient RNA for microarray analysis. As a result, amplified RNA from 1×10^5 SupT1 cells infected with HHV-7 at an apparent MOI of 2.5 was hybridised onto the microarray, which produced good signals, resulting in a 100% pass rate for HHV-7 probes, as compared to 10.2% for unamplified RNA from cells infected at a lower MOI. Therefore a system has been successfully developed for the HHV-7 infection of SupT1 cells that provides sufficient materials for microarray analysis.

CHAPTER 5

GENE EXPRESSION PROFILE OF HHV-7

5.1 Introduction

5.1.1 Classification of herpesvirus gene expression

The expression of herpesvirus genes is tightly regulated during infection (section 1.1.5) and was first classified into immediate early (IE or α), early (E or β) and late (L or γ) by pioneering work on HSV-1 infected cell polypeptides (ICP) in the 1970s, using inhibitors that block *de novo* protein synthesis (Honess and Roizman, 1974) and DNA replication (Honess and Watson, 1977). This classification approach is based on the following observations: 1) the expression of IE genes, which are mainly involved in transcriptional regulation of other viral genes, initiate as soon as the virus enters the host cell or reactivates from latency, and is independent of *de novo* protein synthesis; 2) E genes which encode enzymes and factors for DNA replication, are expressed in the presence of inhibitors of viral DNA replication, but are blocked by inhibitors of *de novo* protein synthesis; 3) expression of structural protein-encoding L genes require both protein synthesis and viral DNA replication, thus are repressed by both types of inhibitors. This approach has been used to classify genes of both human and animal herpesviruses, coupled with various detection methods such as northern blotting, reverse transcription-PCR and microarrays (Table 5.1). One disadvantage of this method, however, is that the effects of inhibitors on the cell may influence the viral transcription program, for example, the inhibitor for viral DNA replication, phosphonoacetic acid (PAA) (at or above 200 μg/ml) completely prevented growth of uninfected cells (Honess and Watson, 1977).

More recently, the use of microarrays for expression profiling without the use of metabolic inhibitors has been applied in the temporal mapping of genes in Kaposi sarcoma-associated herpesvirus (KSHV) (Jenner et al., 2001; Paulose-Murphy et al., 2001) and murine gammaherpesvirus 68 (MHV-68) (Ahn et al., 2002) (Table 5.1). This method relies on the pattern recognition ability of cluster analysis (Section 1.2.3.4) to group together viral genes that are co-regulated at certain stages of the infection. Transcription profiling of KSHV (Jenner et al., 2001) and MHV68 (Ahn et al., 2002) described that genes with similar functions were co-expressed. In KSHV, many DNA replication-related genes were upregulated between 10 – 24 hours post infection (hpi),

while structural genes were co-expressed between 48 – 72 hpi (Jenner et al., 2001). Microarrays had also been coupled with statistical tests to identify viral genes that were differentially up- or downregulated at different stages of the infection (Chambers et al., 1999). The transitions between the different stages of viral infection were sufficiently clear-cut to allow the distinction between genes involved in DNA replication and structural assembly, although some genes may sometimes be detected at the apparently 'wrong' stage of infection. Expression profiling not only allows the classification of genes, but also enables the expression to be monitored in a more authentic infection. In this chapter, the use of expression profiling together with inhibitor-based assays to assign HHV-7 genes into their temporal classes is described.

Table 5.1 Selected publications on temporal mapping of herpesvirus gene expression

Virus [a]	Profiling [b]	Inhibitors	Detection	Reference [c]
Alphaherpesviruses				
HSV-1	N	Y	Radio-labeled protein gels	(Honess and Roizman, 1974)
HSV-2	N	Y	Northern blotting	(Easton and Clements, 1980)
VZV	N	Y	Protein gels	(Lopetegui et al., 1985)
BHV-1	N	Y	Northern blotting	(Seal et al., 1991)
EHV-3	N	Y	Northern blotting	(Sullivan et al., 1990)
Betaherpesviruses				
HCMV	N	Y	Northern blotting	(Wathen and Stinski, 1982)
HCMV	N	Y	Spotted microarrays	(Chambers et al., 1999)
HHV-6	N	Y	Reverse transcription-PCR	(Mirandola et al., 1998)
HHV-7	N	Y	Reverse transcription-PCR	(Menegazzi et al., 1999)
MCMV	N	Y	Northern blotting	(Keil et al., 1984)
Gammaherpesviruses				
KSHV	Y	N	Membrane arrays	(Jenner et al., 2001)
KSHV	Y	N	Spotted microarrays	(Paulose-Murphy et al., 2001)
KSHV	Y	N	Quantitative RT-PCR	(Fakhari and Dittmer, 2002)
MHV-68	Y	Y	Membrane arrays	(Ahn et al., 2002)
MHV-68	Y	Y	Spotted microarrays	(Ebrahimi et al., 2003)
Undefined subfamily				
CCV	N	Y	Northern blotting	(Huang and Hanson, 1998)

[a] HSV-1, Herpes simplex type 1; HSV-2, Herpes simplex type 2; VZV, Varicella-zoster virus; BHV-1, Bovine herpesvirus 1; EHV-3, Equine herpesvirus type 3; HCMV, Human cytomegalovirus; HHV-6, Human herpesvirus type-6; HHV-7, Human herpesvirus type-7; MCMV, Murine cytomegalovirus; KSHV, Kaposi Sarcoma-associated herpesvirus; MHV-68, Murine gammaherpesvirus; CCV, Channel Catfish virus (Ictalurid herpesvirus 1)

[b] Gene expression profiling involving the use of whole transcriptome microarrays or quantitative RT-PCR to profile the change in viral gene expression over time, other studies are qualitative detection of viral transcripts based on metabolic inhibitor assays

c Not a comprehensive list

5.1.2 Temporal regulation of HHV-7 gene expression

The genome of human herpesvirus type-7 (HHV-7) has been sequenced and is predicted to encode 86 ORFs (Megaw et al., 1998; Nicholas, 1996). Functional annotations of these ORFs have been assigned mainly by homology to protein products in other herpesviruses (Table 1.3). Like all herpesviruses, HHV-7 is able to establish a persistent infection in host cells, achieved by the balance between latency and lytic replication (section 1.1.11). A model for studying latent infection of HHV-7 has not been developed, and its lytic replication has only been briefly investigated (Menegazzi et al., 1999). It is believed that lytic replication of HHV-7, like many herpesviruses, proceeds through a temporally regulated transcription programme of lytic gene expression (Roizman, 2001a). The timing of a gene's expression within this programme gives clues to its function (section 1.1.5). For example, genes that are expressed early include enzymes and factors required for DNA replication while those expressed later, after DNA replication has occurred, generally encode structural proteins of the virion.

There has only been one published study on the transcription of HHV-7 genes (Menegazzi et al., 1999), which analysed the expression of 15 ORFs during lytic replication in SupT1 cells using reverse transcription-PCR (RT-PCR). Of these, 10 ORFs (U10, U14, U18, U31, U39, U41, U42, U53, U73, U90) have been categorised as IE genes due to their expression in the presence of inhibitors of *de novo* protein synthesis (cycloheximide, CHX or emetine). U19 and U18/20 transcripts could be detected in the presence of PAA, which blocks viral DNA polymerase, but not in the presence of emetine, hence was classified as E genes. The unspliced mRNAs of U17 and U66 were IE transcripts, whereas their spliced versions were E and L transcripts, respectively. U100 was shown to have multiple splice sites, giving both IE and L transcripts when spliced and the unspliced mRNA showed E transcription kinetics.

Of HHV-7's 86 predicted genes, the expression of less than 20% has been analysed by only one study. Therefore, the expression pattern of many HHV-7 genes, and how this relates to the ordered series of events during lytic replication, remain largely unknown. The attention given to this virus is far less than the other members of the

betaherpesvirus subfamily, human herpesvirus type-6 (HHV-6) and human cytomegalovirus (HCMV). This is mostly due to the lack of confirmed disease association for HHV-7. Nevertheless, HHV-7 is an interesting candidate for study because of its low pathogenicity. This virus infects more than 90% of the world's population and persists in the host without apparent ill effects (section 1.1.8). We still do not know whether it really exists in a commensal symbiosis with human, or whether it is an opportunistic pathogen. HHV-7 reactivates HHV-6 from latency hence may contribute to complications in the immunocompromised (section 1.1.9). Alternatively, the virus may actually benefit the host in certain instances by competing for the CD4 receptor on T-cells with human immunodeficiency virus-1 (HIV-1) (Lusso et al., 1994), and downregulating surface CD4 to prevent super-infection by HIV-1 (section 1.1.10).

To gain more understanding about HHV-7 infection of T-cells, the viral lytic replication programme could be reconstructed by the large scale measurement of viral transcription using microarray technology (section 1.2.2), coupled with the pattern identification using cluster analysis (section 1.2.3.4). This chapter aims to use the custom-made HHV-7 microarray (section 3) to analyse HHV-7 gene expression during lytic replication in SupT1 cells.

5.2 Results

5.2.1 HHV-7 infection of SupT1 cells

Custom-made host-pathogen microarrays were used to monitor changes in the expression of HHV-7 genes during *in vitro* infection of SupT1 cells (a $CD4^+$ T-cell line) over a 72-hour period. Cells were infected with cell-free HHV-7 virus at a multiplicity of infection (MOI) of approximately 2.5, with spinoculation at 1000 g. Cells were harvested at each time point: 8, 12, 16, 20, 24, 38, 48, 72 hours post infection (hpi) for RNA extraction and analysis by microarrays. The level of viral infection was confirmed by immunofluorescence assay (IFA) of selected samples using an anti-HHV-7 monoclonal antibody (Figure 5.1). Percentage of HHV-7-specific antigen-bearing cells increased with time, peaking at 48 hpi with approximately 80% of the cells infected (Figure 5.2b). The percentage of infected cells declined between 48 and 72 hpi, probably due to the proliferation of uninfected cells. The large percentage of infection was important for the microarray analysis, as this meant that the gene expression patterns detected would be representative of HHV-7 infected cells, as opposed to uninfected ones. The virus inoculum was removed after spinoculation, hence the infection should be synchronous (i.e. infection of all cells initiated within the first hour). However, it might be possible that virus removal was not 100% efficient and a smaller fraction of cells were infected at later times, causing asynchronous infection, which may explain the increase in the percentage of infected cells over time (up to 48 hpi). It was also possible that the infected cells began expressing HHV-7 antigen at different times, due to different cells being at different stages of infection. Nonetheless, the effects caused by this were minimal in the gene expresson analysis results (section 5.2.3), which measured the relative rather than absolute abundance of genes across time.

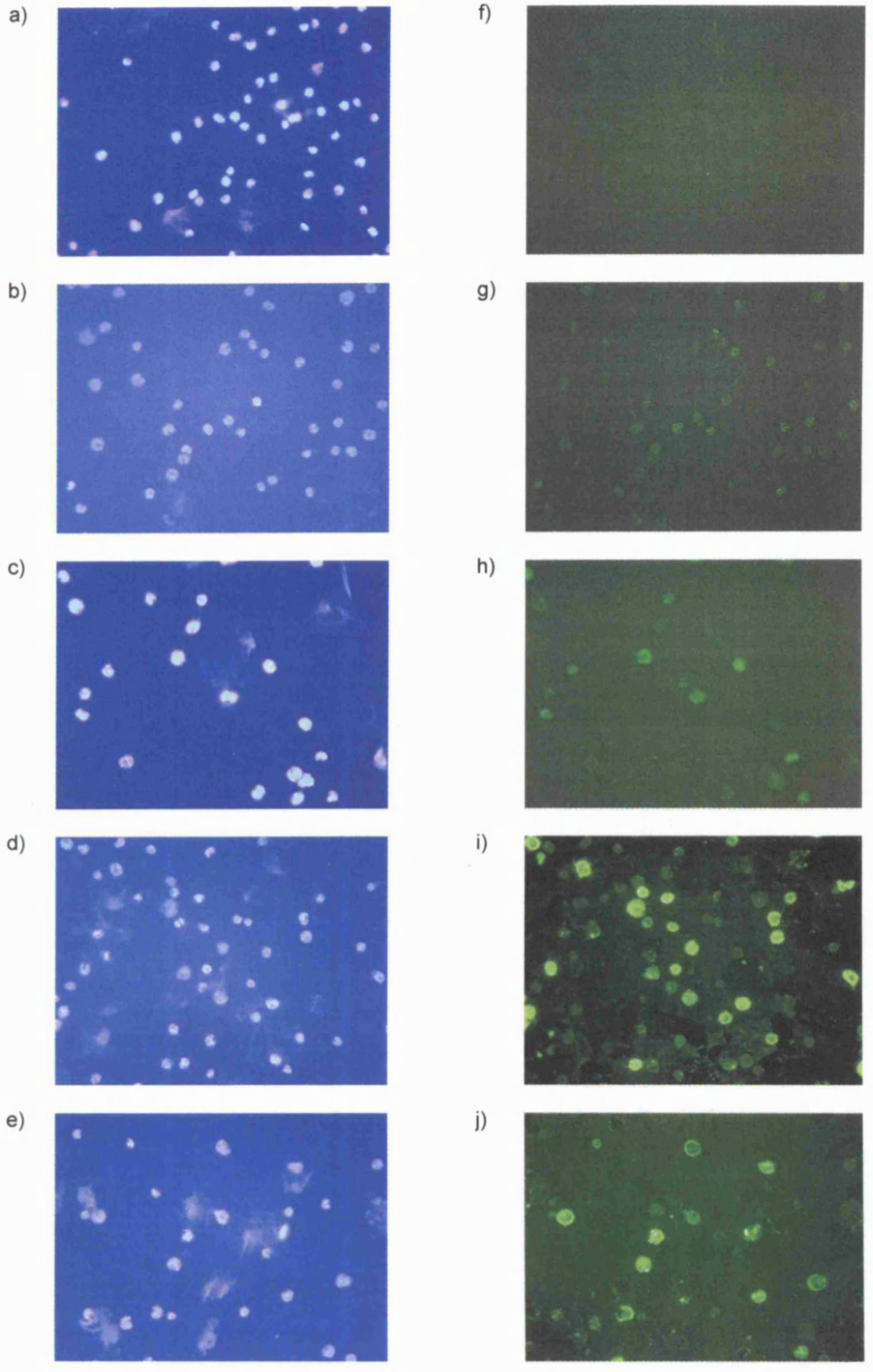


Figure 5.1 Immunofluorescence assay of HHV-7 infected Sup-T1 cells. Mock-infected (a, f); or HHV-7-infected cells harvested at 8 hours post infection (b, g); 24 hpi (c, h); 48 hpi (d, i); 72 hpi (e, j) were fixed on slides and infection detected using primary antibody targeting HHV-7 antigen. Immunofluorescence was visualised using FITC-conjugated secondary antibody (f-j). DAPI counter-stain showed total number of cells in each corresponding fields of view (a-e).

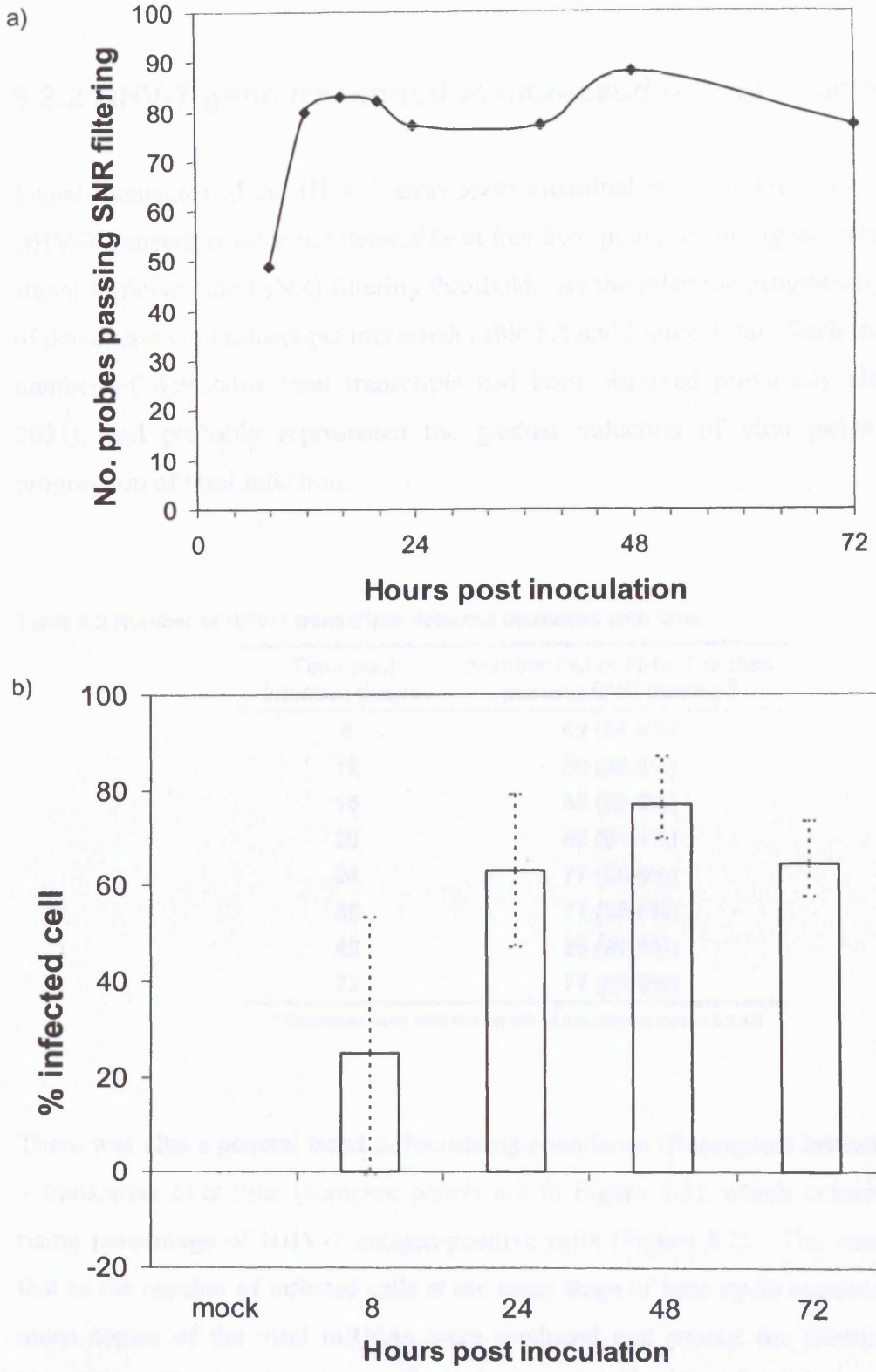


Figure 5.2 a) Number of HHV-7 probes passing the SNR filtering criteria during time course; b) percentage of infected cells at selected time points. Percentage of infected cells was monitored by immunofluorescence assay and was calculated as: (number of green fluorescent cells / total number of cells) × 100, results shown are the average of four fields of view per time point. Dotted lines indicate standard deviation.

5.2.2 HHV-7 gene transcription increased during infection

Signal intensities of the HHV-7 array spots measured at 8 hpi were low. Half of the HHV-7 transcripts were not detectable at this time point, as the signals were below the signal-to-noise ratio (SNR) filtering threshold. As the infection progressed, the number of detectable viral transcripts increased(Table 5.2 and Figure 5.2a). Such increase in the number of detectable viral transcripts had been observed previously (Jenner et al., 2001), and probably represented the gradual induction of viral genes during the progression of viral infection.

Table 5.2 Number of HHV-7 transcripts detected increased with time.

Time post-infection (hours)	Number (%) of HHV-7 probes passing SNR filtering [a]
8	49 (54.4%)
12	80 (88.9%)
16	83 (92.2%)
20	82 (91.1%)
24	77 (85.6%)
38	77 (85.6%)
48	88 (97.8%)
72	77 (85.6%)

[a] Determined using SNR filtering criteria described in section 3.2.4.2

There was also a general trend of increasing abundance (fluorescent intensity) of HHV-7 transcripts over time (compare panels a-h in Figure 5.3), which coincided with the rising percentage of HHV-7 antigen-positive cells (Figure 5.2). The reason could be that as the number of infected cells at the same stage of lytic cycle expanded over time, more copies of the viral mRNAs were produced and passed the filtering threshold. However, although the intensities gave a general picture of the trend in viral transcription, intensities of individual genes cannot be directly compared between different samples. Comparisons between two-colour arrays need to be made using the Cy5/Cy3 ratios (explained in section 1.2.2), and will be described in later sections. However, the intensities of different genes can be compared within an array, showing that some gene probes had higher signals than others. The intensities of probes U27,

U36, U42, U81, U90 and X5 were high relative to the other probes in almost every array; while signals of probes U53.5 and U56 were prominent in all arrays from 38hpi onwards. Of these, probes U36, U42 and X5 have been shown to be of slightly lower specificity, producing just above-background signals on arrays hybridised to mock-infected samples (section 3.2.4.2), which may explain their higher-than-average signals. The intense signals might also be the result of higher hybridisation efficiency of these array probes; or perhaps these genes were indeed expressed at higher levels. The HCMV homologue of U90 (HCMV UL122/3) and U27 (UL44) are known to be expressed in high abundance (Chambers et al., 1999; Geballe et al., 1986; Gibson, 1983; Stinski, 1978).

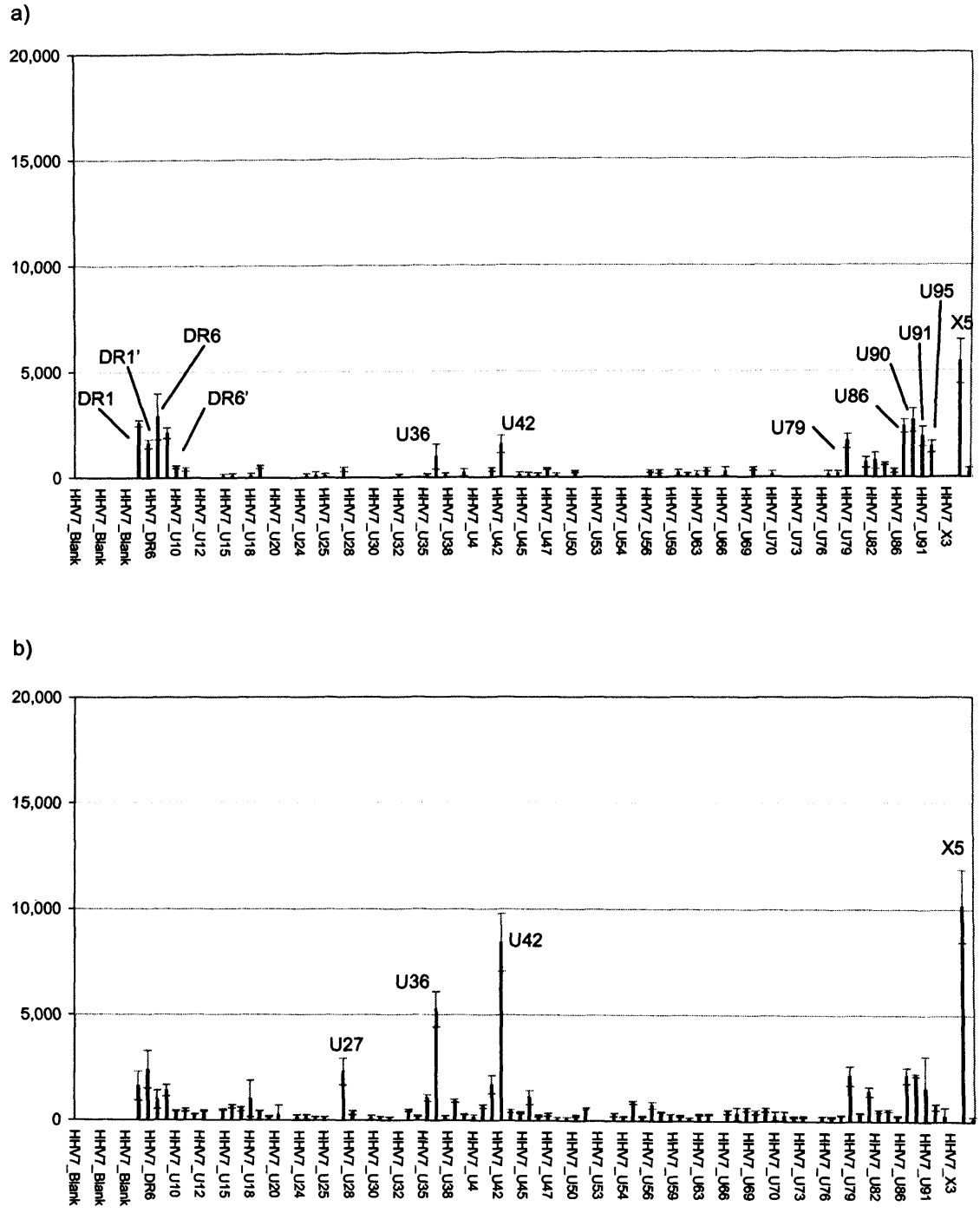
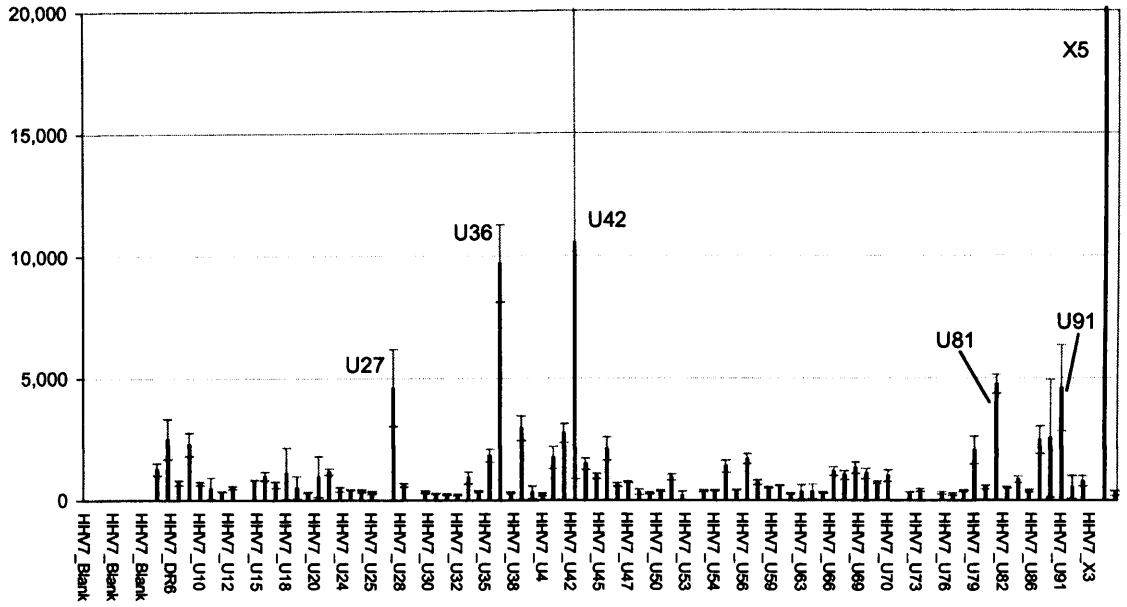


Figure 5.3 Cy5 signal intensities of HHV-7 array spots on microarrays hybridised to labelled cDNA from a) 8 hpi, b) 12 hpi, c) 16 hpi, d) 20 hpi, e) 24 hpi, f) 38 hpi, g) 48 hpi, h) 72 hpi. Data are the average of triplicate spots on each array; have been background-subtracted, filtered and subjected to total intensity normalisation and Lowess normalisation.

(c)



(d)

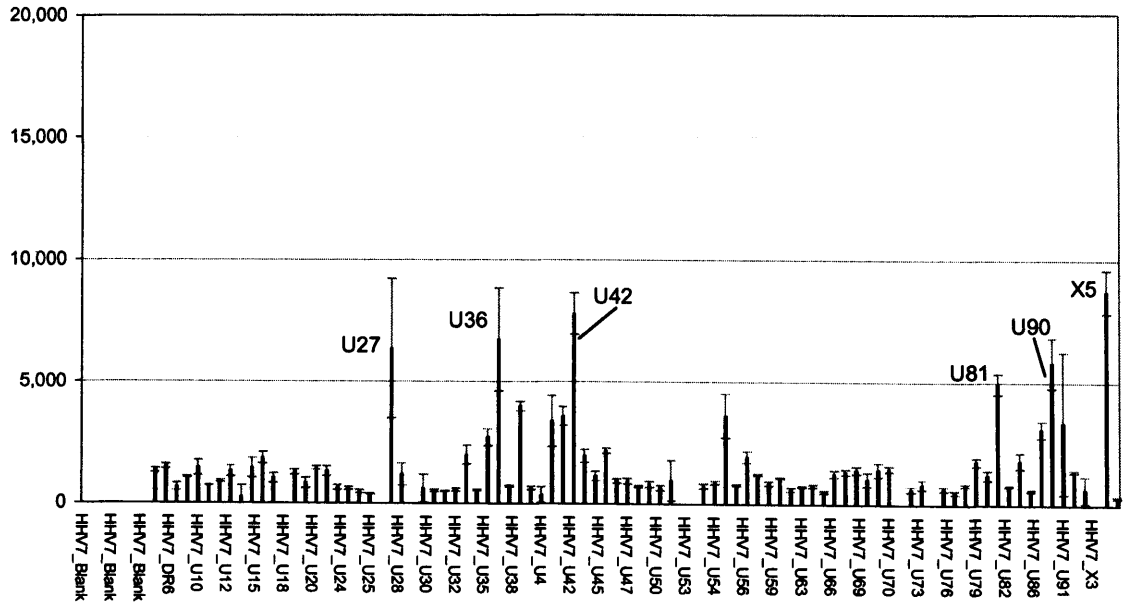
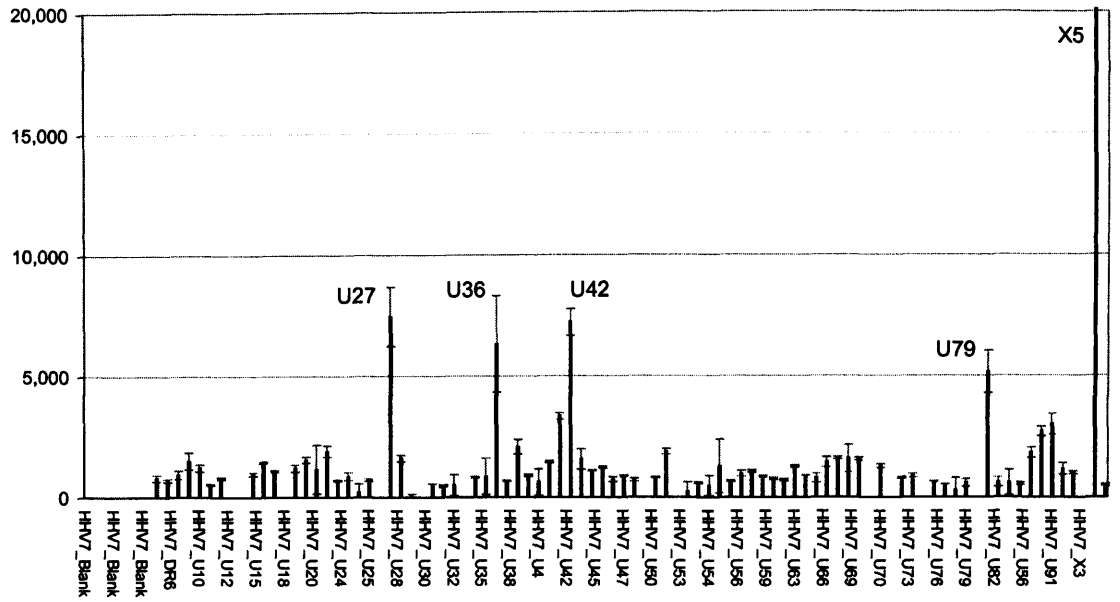


Figure 5.3 (cont.)

(e)



(f)

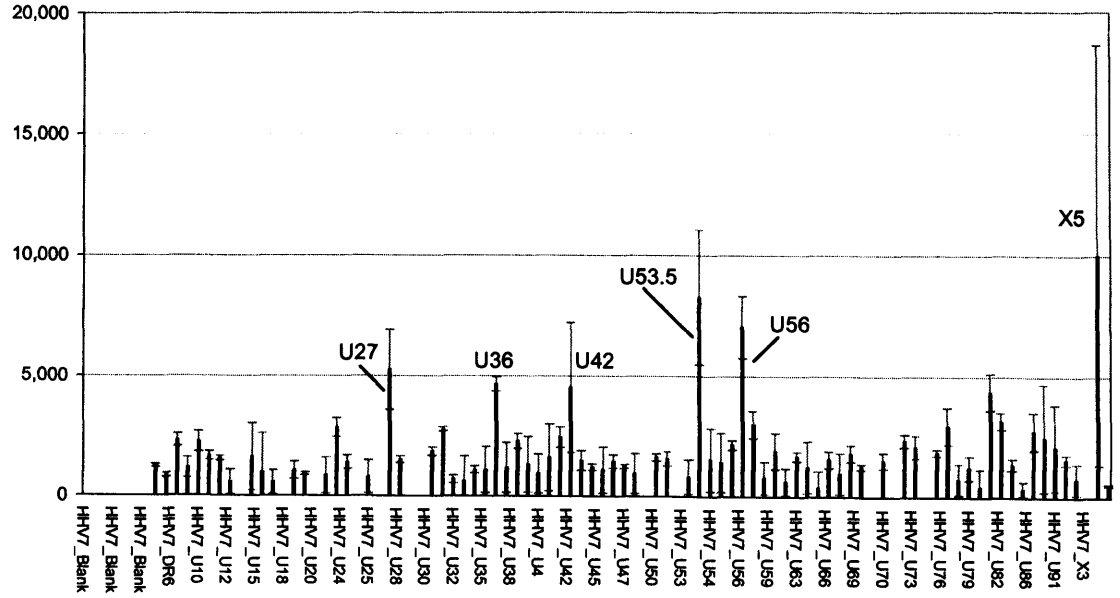
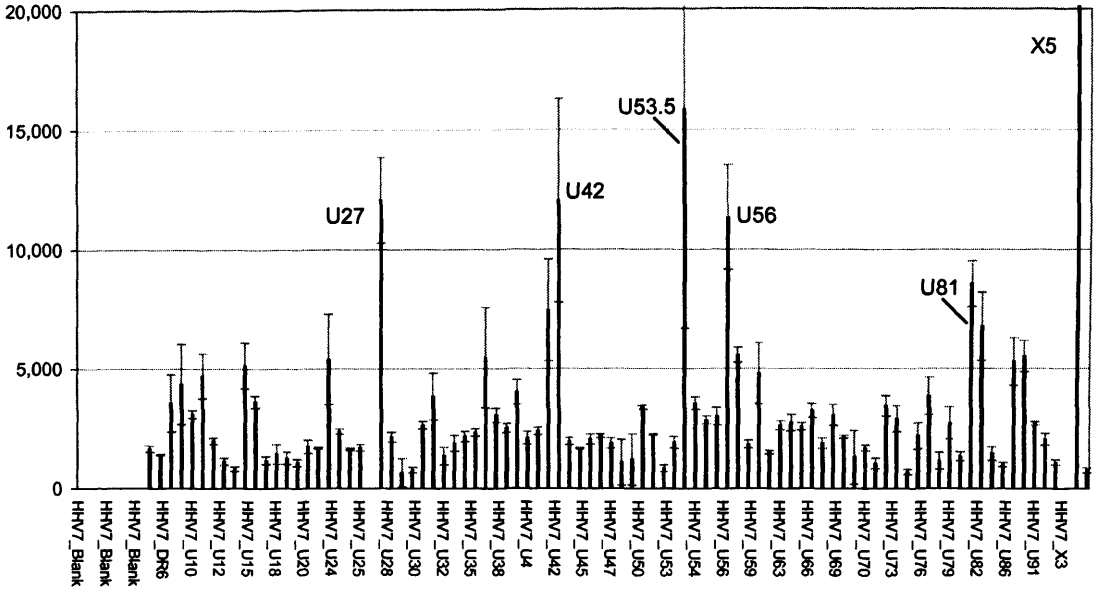


Figure 5.3 (cont.)

(g)



(h)

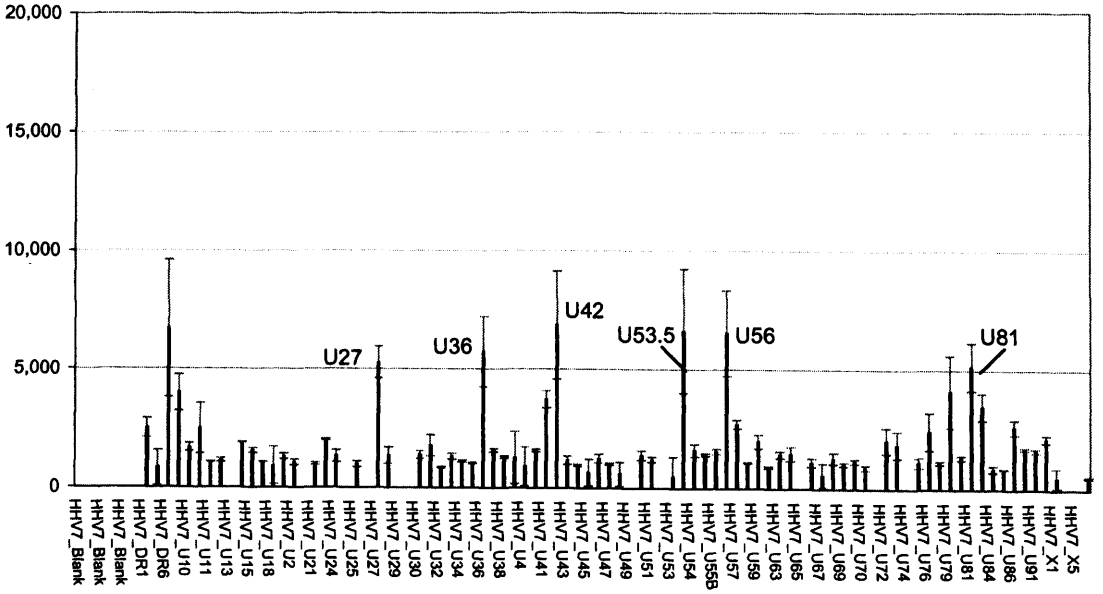


Figure 5.3 (cont.)

5.2.3 HHV-7 expression profile during lytic infection of SupT1 cells

After log_2 transformation, normalisation and mean-centring (section 1.2.3) to enable cross-array comparison, microarray data were analysed by hierarchical clustering to arrange time points (from left to right) and genes (from top to bottom) according to similarity of the gene expression patterns (Figure 5.4 and 5.5). Two cluster analyses were performed, the first one (100%-presence) was with more stringent criteria, including only array spots that passed filtering (section 3.2.4.2) in all arrays (Figure 5.4). The result was therefore unaffected by missing data, although a fewer number of genes were included. The second cluster analysis (80%-presence) included genes with missing values in up to 1 out of 8 arrays. This may affect the stability of clustering pattern (de Brevern et al., 2004), but allowed more genes to be analysed (Figure 5.5).

5.2.3.1 Clustering (100%-presence) of HHV-7 genes

HHV-7 gene expression during a 72-hour time course infection of SupT1 cells were analysed by hierarchical clustering, which arranged the 8 time points into three main clusters from left to right (Figure 5.4): i) 8-12 hpi, ii) 16-24 hpi, and iii) 38-72 hpi. Within the 16-24 hpi cluster, the expression profiles at 16 and 20 hpi were found to be most similar to each other, whereas 24 hpi had the most distant profile. Profiles of 38 hpi and 48 hpi were more similar to each other than to that of 72 hpi. The ordering of the 8 time points on the cluster therefore reflected a gradual change in HHV-7 gene expression throughout the progression of infection. The largest shift in overall gene expression occurred between 24 to 38 hpi, and the second largest from 12 to 16 hpi, indicating major changes in viral gene expression between these times, which will be discussed in more detail later.

Hierarchical clustering also arranged the genes into three main clusters (I, II and III) from top to bottom, some of which can be viewed as two sub-clusters (IIA and IIB; IIIA and IIIB) (Figure 5.4). The mean gene expression profiles of the different clusters were also visualised in graphical form (Figure 5.6). Genes in the top cluster (I) mainly

showed decreased expression over time, followed by an increase between 48 and 72 hpi (Figure 5.6a). The middle cluster (II) was made up of two sub-clusters, the first of which (IIA) consisted of genes that showed early (8 – 12 hpi) increase in expression, followed by a decline (Figure 5.6b). Genes in second sub-cluster (IIB) had higher expression levels at 8hpi (Figure 5.6c), therefore were probably induced earlier in infection than those in IIA. Genes in cluster IIIA had a similar profile to IIA, except that the major rise in expression took place slightly later (between 12 and 16 hpi) and the decline after 24 hpi was not as rapid (Figure 5.6d). Finally, the bottom cluster (IIIB) consisted of genes whose expressions increased more slowly, and at a later stage (24 and 38 hpi) in infection (Figure 5.6e).

Although each gene probe was spotted in triplicate on the array, some genes with lower abundance only appeared once or twice on the cluster diagram, because one or two spots have failed the filtering criteria. However, the replicates were closely located in the cluster, indicating the replicate probes gave reproducible results, detecting similar expression profiles for the same genes.

Figure 5.4 (next page)
Cluster analysis of HHV-7 gene expression over a 72-hour infection of SupT1 cells (MOI=2.5). Only 80 viral gene spots (37 ORFs) with no missing data were included in analysis to give this 100%-presence cluster diagram. Time point samples (i.e. array slides) were clustered horizontally; viral genes were clustered vertically into three major cluster groups I, II and III. A value between -3.0 to +3.0 was calculated for each gene at each time point, depending on the relative level of gene expression. The values were presented on a colour scale (-3.0, green, downregulated; 0.0, black, unchanged; +3.0, red, upregulated). Gene names were colour-coded by their predicted function: green, transactivators; blue, DNA replication; orange, capsid & assembly; pink, tegument; red, envelope, grey, others; black, unknown.

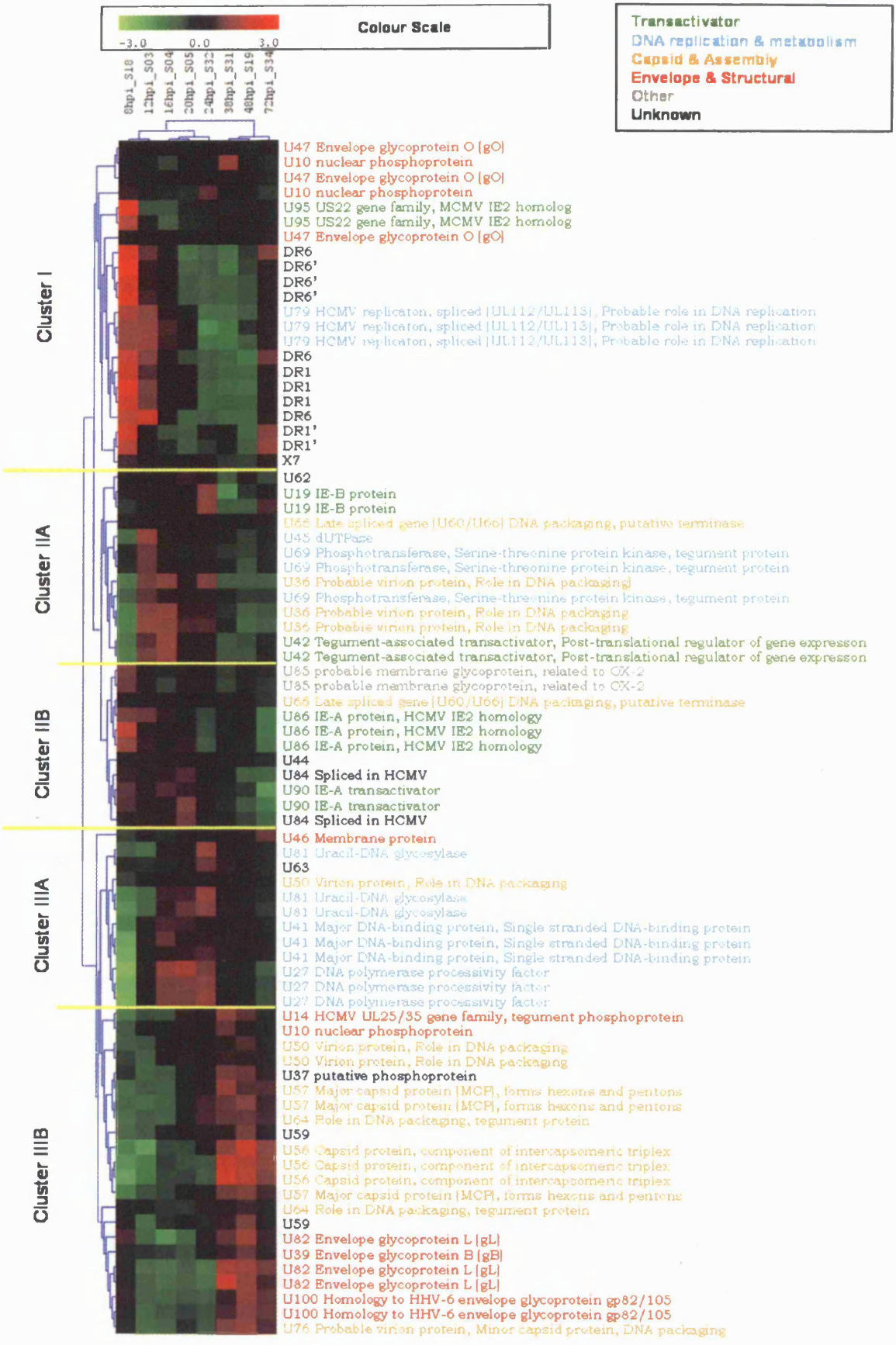


Figure 5.4 Cluster analysis results (100%-presence)

5.2.3.2 Clustering (80%-presence) of HHV-7 genes

The clustering analyses in section 5.2.3.1 included only array data points that passed the data filtering criteria across all 8 arrays, therefore only 37 ORFs were included. In order to analyse the expression of more viral genes, the clustering parameters were amended so that gene spots that passed filtering in more than 80% of arrays (7 out of 8 arrays) could be included in clustering analysis (Figure 5.5). This adjustment allowed 31 additional genes to be clustered, therefore a total of 68 out of 90 (76%) HHV-7 probes were analysed by hierarchical clustering. The remaining 24% were excluded from clustering analysis due to low signals over the entire time course, which may be the result of probe design or low abundance transcripts.

The overall clustering pattern in this 80%-presence clustering (Figure 5.5a) was slightly different to that of the 100%-presence clustering (Figure 5.4), but the majority of genes were assigned to the same sub-clusters as before. The reason for this slight difference was probably because as the clusters were formed by the agglomerative average-linkage method, the inclusion of the extra 31 genes altered the new distances calculated between clusters (section 1.2.3.4), resulting in the sub-clusters being positioned differently. The expression profile at 12 hpi was now more similar to 16 hpi than to 8 hpi as suggested in the 100%-presence cluster. This indicated the gene expression pattern actually changed gradually from 8 through 12 to 16 hpi, unlike the major shift between 24 and 38 hpi, which was observed in both 100%- and 80%-presence clusters. More changes were observed in the arrangement of gene clusters. Cluster IIA was now more similar to clusters IIIA and I than to IIB; whereas cluster IIIB was the most distant. This illustrated the characteristic of the agglomerative nature of average-linkage hierarchical clustering (section 1.2.3.4). In light of this, it was important not to focus on the top-to-bottom transitions, but to interpret both 100%- and 80%-presence clustering results in the form of the smaller and more robust sub-clusters. In the rare cases where genes were allocated to different sub-clusters by 100% and 80% clustering (for example U95), the assignment in the 100% cluster (cluster I) was considered to be more accurate.

Figure 5.5 (next page)
a) Cluster analysis of HHV-7 gene expression over a 72h-infection of SupT1 cells (MOI=2.5). 168 viral gene spots (68 ORFs), with missing values (grey squares) in up to one time point sample were analysed to give this 80%-presence cluster diagram. Time point samples were clustered horizontally; viral genes were clustered vertically. Expression data at 0 hpi was not included in cluster analysis but presented on the side for comparison. b) Zoomed image of Cluster groups I, IIIA, IIA; c) Zoomed image of Cluster group IIB; d) Zoomed image of Cluster group IIIB. See also legend for Figure 5.4

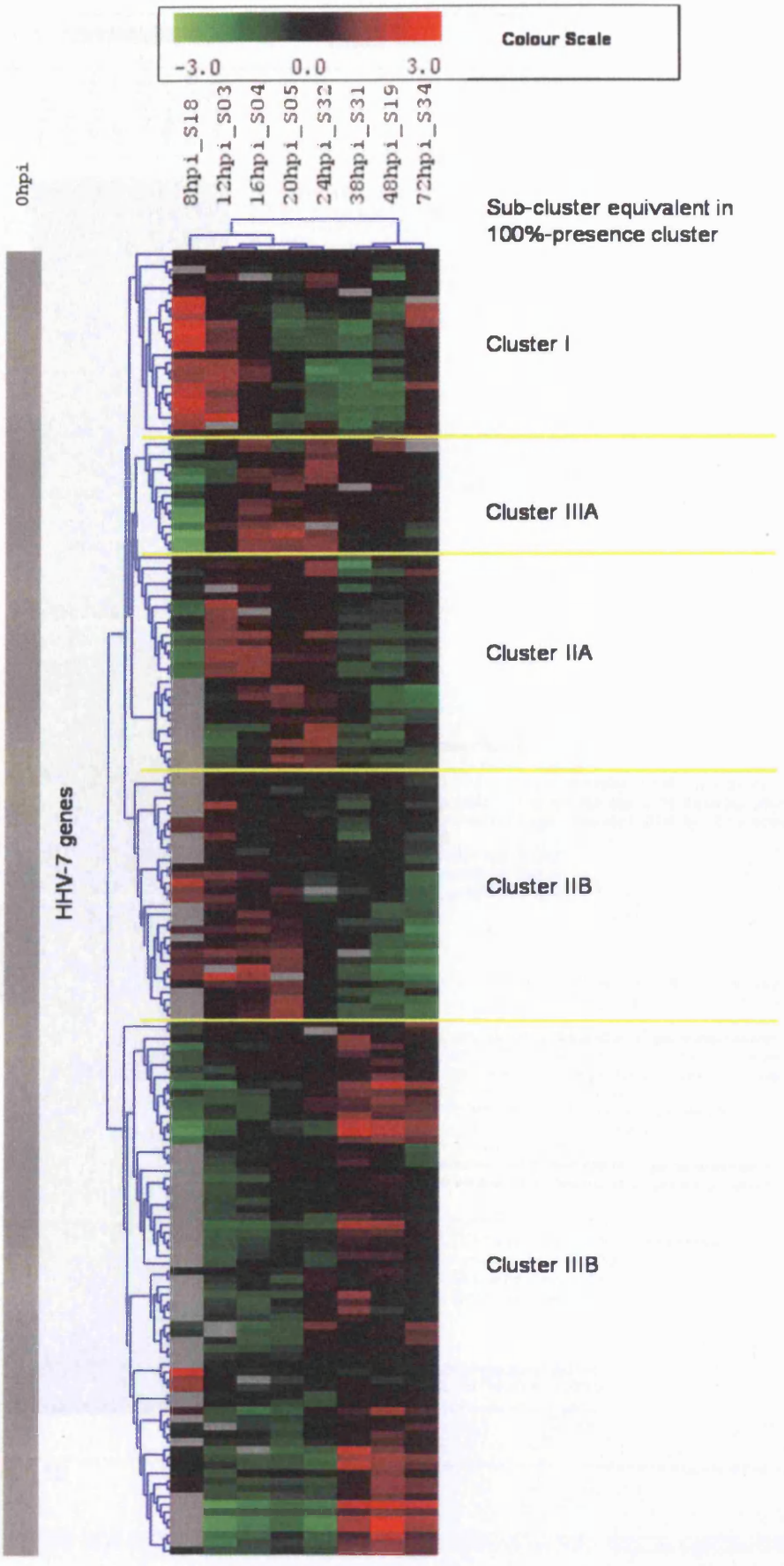


Figure 5.5a Clustering results (80%-presence)

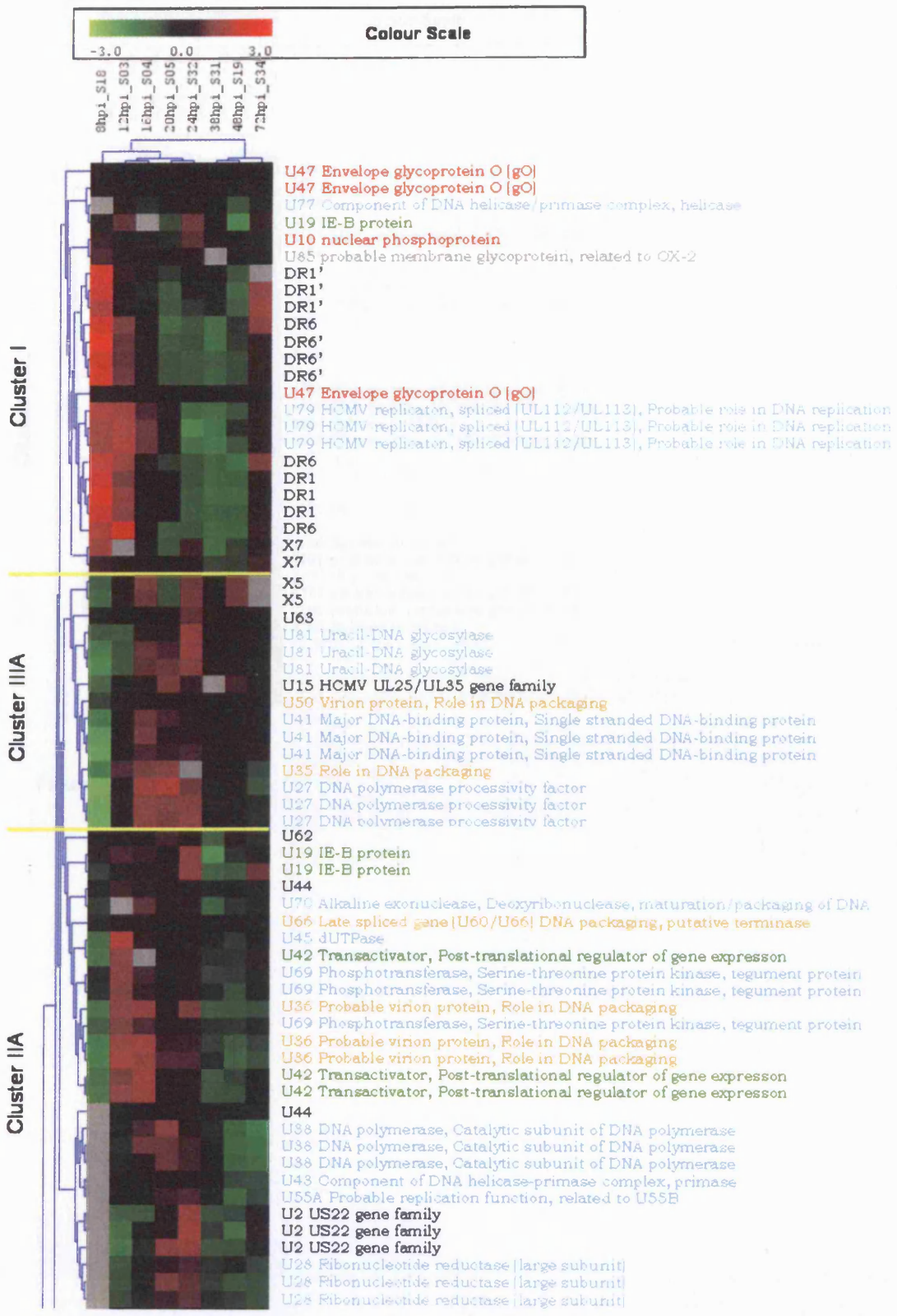

Colour Scale
-3.0
0.0
3.0
8hpi_S18
12hpi_S03
16hpi_S04
20hpi_S05
24hpi_S32
38hpi_S31
48hpi_S19
72hpi_S34
Cluster I
Cluster IIIA
Cluster IIA
U47 Envelope glycoprotein O [gO]
U47 Envelope glycoprotein O [gO]
U77 Component of DNA helicase/primase complex, helicase
U19 IE-B protein
U10 nuclear phosphoprotein
U85 probable membrane glycoprotein, related to OX-2
DR1'
DR1'
DR1'
DR6
DR6'
DR6'
DR6'
U47 Envelope glycoprotein O [gO]
U79 HCMV replicaton, spliced [UL112/UL113], Probable role in DNA replication
U79 HCMV replicaton, spliced [UL112/UL113], Probable role in DNA replication
U79 HCMV replicaton, spliced [UL112/UL113], Probable role in DNA replication
DR6
DR1
DR1
DR1
DR6
X7
X7
X5
X5
U63
U81 Uracil-DNA glycosylase
U81 Uracil-DNA glycosylase
U81 Uracil-DNA glycosylase
U15 HCMV UL25/UL35 gene family
U50 Virion protein, Role in DNA packaging
U41 Major DNA-binding protein, Single stranded DNA-binding protein
U41 Major DNA-binding protein, Single stranded DNA-binding protein
U41 Major DNA-binding protein, Single stranded DNA-binding protein
U35 Role in DNA packaging
U27 DNA polymerase processivity factor
U27 DNA polymerase processivity factor
U27 DNA polymerase processivity factor
U62
U19 IE-B protein
U19 IE-B protein
U44
U70 Alkaline exonuclease, Deoxyribonuclease, maturation/packaging of DNA
U66 Late spliced gene [U60/U66] DNA packaging, putative terminase
U45 dUTPase
U42 Transactivator, Post-translational regulator of gene expresson
U69 Phosphotransferase, Serine-threonine protein kinase, tegument protein
U69 Phosphotransferase, Serine-threonine protein kinase, tegument protein
U36 Probable virion protein, Role in DNA packaging
U69 Phosphotransferase, Serine-threonine protein kinase, tegument protein
U36 Probable virion protein, Role in DNA packaging
U36 Probable virion protein, Role in DNA packaging
U42 Transactivator, Post-translational regulator of gene expresson
U42 Transactivator, Post-translational regulator of gene expresson
U44
U38 DNA polymerase, Catalytic subunit of DNA polymerase
U38 DNA polymerase, Catalytic subunit of DNA polymerase
U38 DNA polymerase, Catalytic subunit of DNA polymerase
U43 Component of DNA helicase-primase complex, primase
U55A Probable replication function, related to U55B
U2 US22 gene family
U2 US22 gene family
U2 US22 gene family
U28 Ribonucleotide reductase [large subunit]
U28 Ribonucleotide reductase [large subunit]
U28 Ribonucleotide reductase [large subunit]

Figure 5.5b

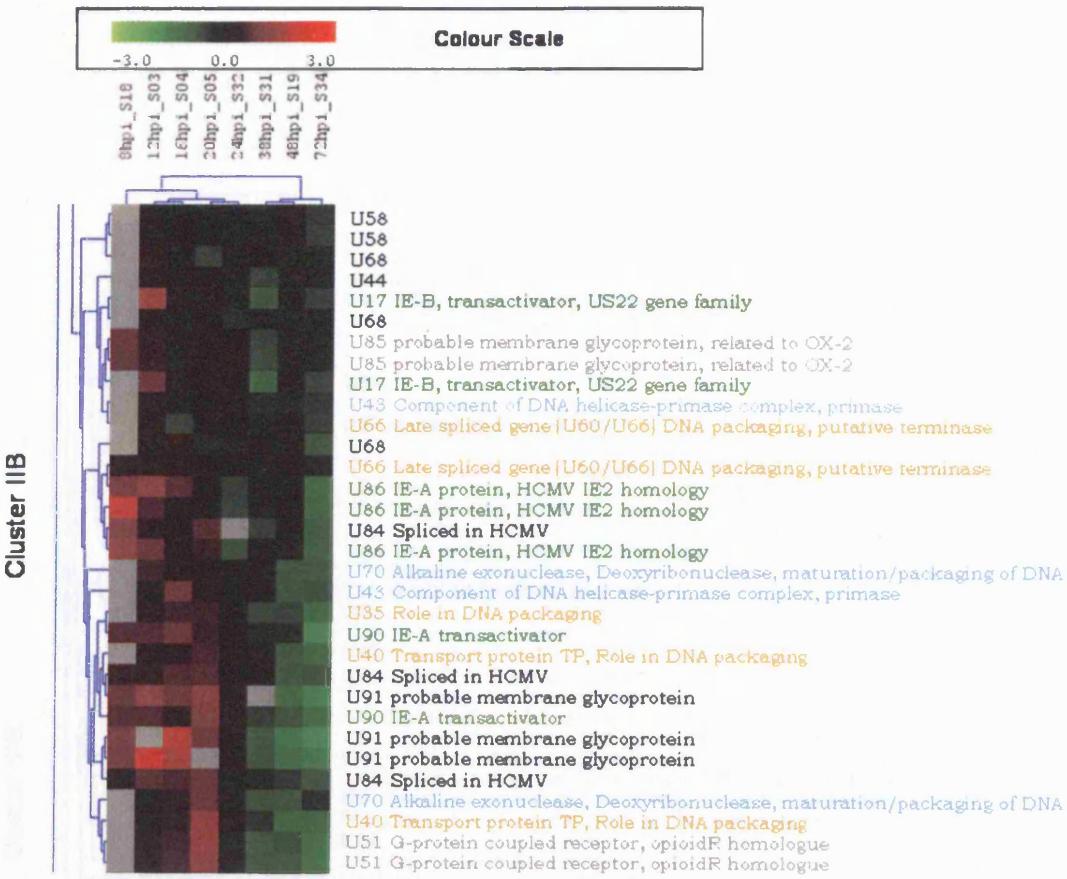

Colour Scale
-3.0
0.0
3.0
8hpi_S18
12hpi_S03
16hpi_S04
20hpi_S05
24hpi_S32
36hpi_S31
48hpi_S19
72hpi_S34
Cluster IIB
U58
U58
U68
U44
U17 IE-B, transactivator, US22 gene family
U68
U85 probable membrane glycoprotein, related to OX-2
U85 probable membrane glycoprotein, related to OX-2
U17 IE-B, transactivator, US22 gene family
U43 Component of DNA helicase-primase complex, primase
U66 Late spliced gene [U60/U66] DNA packaging, putative terminase
U68
U66 Late spliced gene [U60/U66] DNA packaging, putative terminase
U86 IE-A protein, HCMV IE2 homology
U86 IE-A protein, HCMV IE2 homology
U84 Spliced in HCMV
U86 IE-A protein, HCMV IE2 homology
U70 Alkaline exonuclease, Deoxyribonuclease, maturation/packaging of DNA
U43 Component of DNA helicase-primase complex, primase
U35 Role in DNA packaging
U90 IE-A transactivator
U40 Transport protein TP, Role in DNA packaging
U84 Spliced in HCMV
U91 probable membrane glycoprotein
U90 IE-A transactivator
U91 probable membrane glycoprotein
U91 probable membrane glycoprotein
U84 Spliced in HCMV
U70 Alkaline exonuclease, Deoxyribonuclease, maturation/packaging of DNA
U40 Transport protein TP, Role in DNA packaging
U51 G-protein coupled receptor, opioidR homologue
U51 G-protein coupled receptor, opioidR homologue

Figure 5.5c

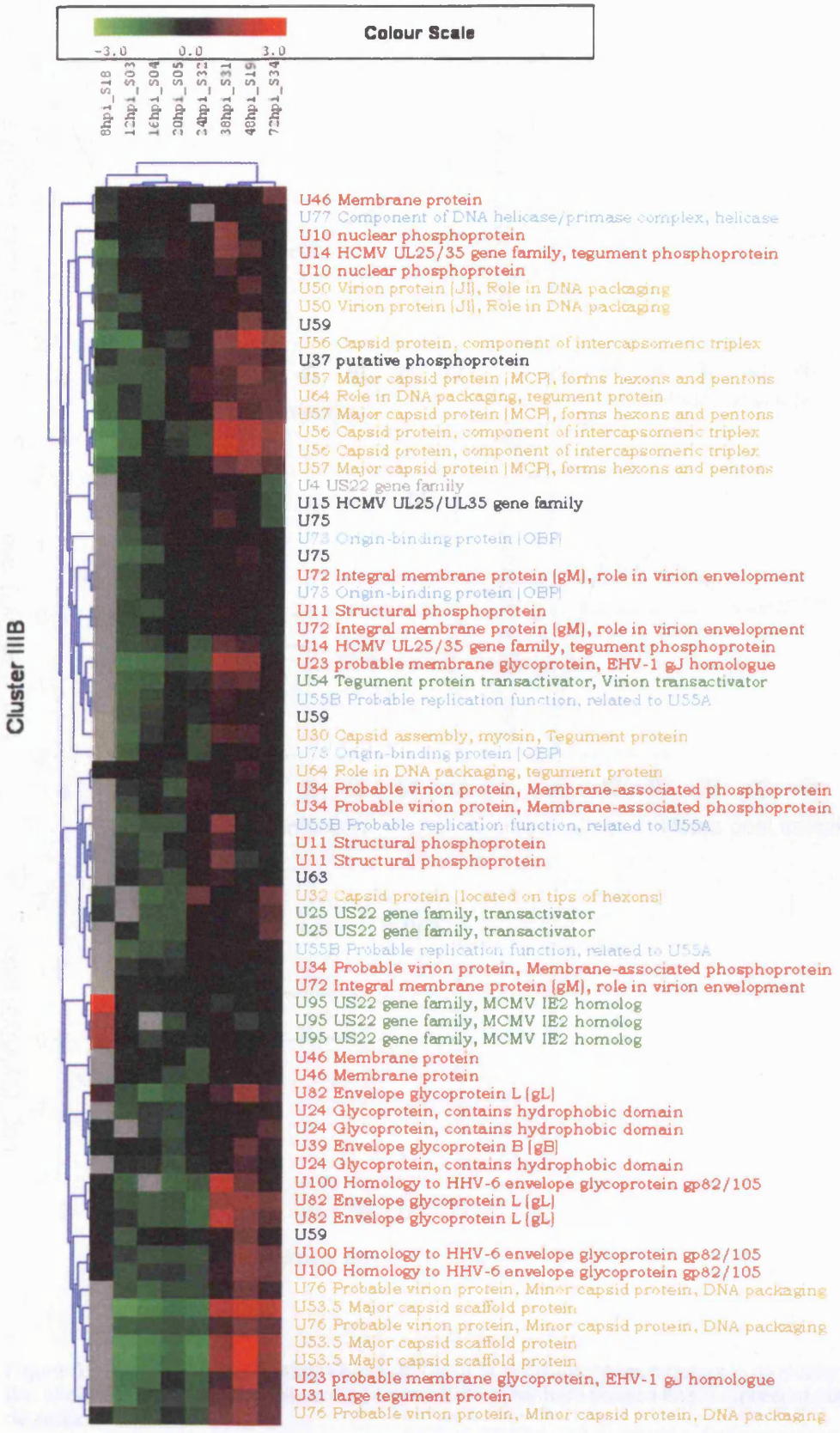

Colour Scale
-3.0
0.0
3.0
8hpi_S18
12hpi_S03
16hpi_S04
20hpi_S05
24hpi_S32
38hpi_S31
48hpi_S19
72hpi_S34
Cluster IIIB
U46 Membrane protein
U77 Component of DNA helicase/primase complex, helicase
U10 nuclear phosphoprotein
U14 HCMV UL25/35 gene family, tegument phosphoprotein
U10 nuclear phosphoprotein
U50 Virion protein [J1], Role in DNA packaging
U50 Virion protein [J1], Role in DNA packaging
U59
U56 Capsid protein, component of intercapsomeric triplex
U37 putative phosphoprotein
U57 Major capsid protein [MCP], forms hexons and pentons
U64 Role in DNA packaging, tegument protein
U57 Major capsid protein [MCP], forms hexons and pentons
U56 Capsid protein, component of intercapsomeric triplex
U56 Capsid protein, component of intercapsomeric triplex
U57 Major capsid protein [MCP], forms hexons and pentons
U4 US22 gene family
U15 HCMV UL25/UL35 gene family
U75
U73 Origin-binding protein [OBP]
U75
U72 Integral membrane protein [gM], role in virion envelopment
U73 Origin-binding protein [OBP]
U11 Structural phosphoprotein
U72 Integral membrane protein [gM], role in virion envelopment
U14 HCMV UL25/35 gene family, tegument phosphoprotein
U23 probable membrane glycoprotein, EHV-1 gJ homologue
U54 Tegument protein transactivator, Virion transactivator
U55B Probable replication function, related to U55A
U59
U30 Capsid assembly, myosin, Tegument protein
U73 Origin-binding protein [OBP]
U64 Role in DNA packaging, tegument protein
U34 Probable virion protein, Membrane-associated phosphoprotein
U34 Probable virion protein, Membrane-associated phosphoprotein
U55B Probable replication function, related to U55A
U11 Structural phosphoprotein
U11 Structural phosphoprotein
U63
U32 Capsid protein [located on tips of hexons]
U25 US22 gene family, transactivator
U25 US22 gene family, transactivator
U55B Probable replication function, related to U55A
U34 Probable virion protein, Membrane-associated phosphoprotein
U72 Integral membrane protein [gM], role in virion envelopment
U95 US22 gene family, MCMV IE2 homolog
U95 US22 gene family, MCMV IE2 homolog
U95 US22 gene family, MCMV IE2 homolog
U46 Membrane protein
U46 Membrane protein
U82 Envelope glycoprotein L [gL]
U24 Glycoprotein, contains hydrophobic domain
U24 Glycoprotein, contains hydrophobic domain
U39 Envelope glycoprotein B [gB]
U24 Glycoprotein, contains hydrophobic domain
U100 Homology to HHV-6 envelope glycoprotein gp82/105
U82 Envelope glycoprotein L [gL]
U82 Envelope glycoprotein L [gL]
U59
U100 Homology to HHV-6 envelope glycoprotein gp82/105
U100 Homology to HHV-6 envelope glycoprotein gp82/105
U76 Probable virion protein, Minor capsid protein, DNA packaging
U53.5 Major capsid scaffold protein
U76 Probable virion protein, Minor capsid protein, DNA packaging
U53.5 Major capsid scaffold protein
U53.5 Major capsid scaffold protein
U23 probable membrane glycoprotein, EHV-1 gJ homologue
U31 large tegument protein
U76 Probable virion protein, Minor capsid protein, DNA packaging

Figure 5.5d

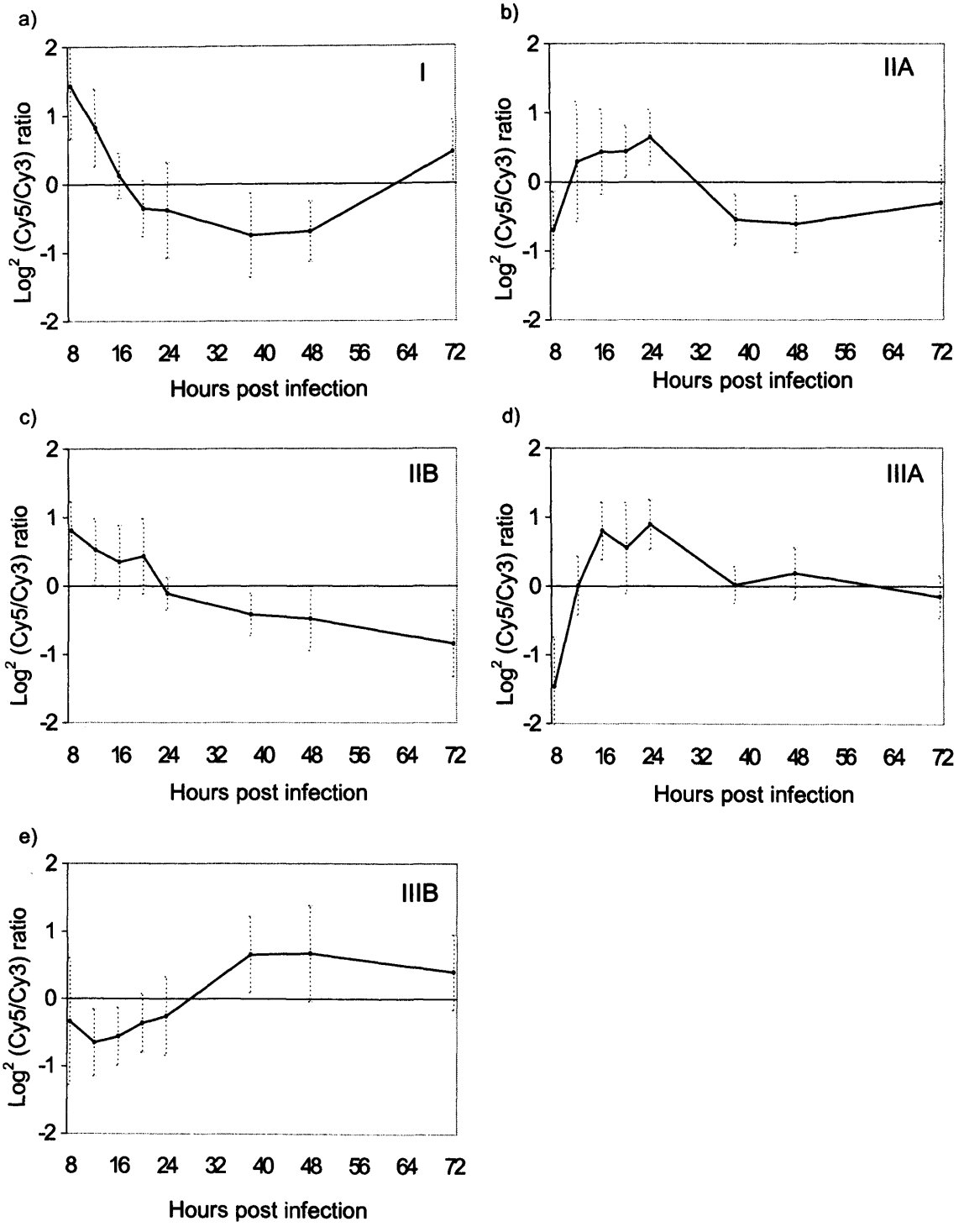


Figure 5.6 Mean expression ratios [$\log_2$(Cy5/Cy3)] of 100%-prescence genes in a) cluster I; b) cluster IIA; c) cluster IIB; d) cluster IIIA; and e) cluster IIIB. Error bars (dotted lines) represent standard deviation.

5.2.3.3 HHV-7 genes of similar functions were co-expressed

HHV-7 genes were grouped into 4 functional classes by their predicted functions. These were transcription regulation, DNA replication & metabolism, DNA packaging & capsid formation, and virion assembly. These were colour-coded on the cluster diagram (Figure 5.4 & 5.5), together with genes of other functions that did not belong to these four groups, and genes of unknown function. Often genes with similar functions were clustered together. The majority of transcription regulators were assigned into cluster IIB; genes involved in DNA replication into clusters IIA and IIIA; structural genes for virion assembly into cluster IIIB. The most likely explanation for this clustering pattern was because viral genes with similar functions were involved in the same stage of the infection cycle, hence co-expressed at certain times during infection (Figure 5.7). This finding was consistent with the expression profiling of KSHV (Jenner et al., 2001; Paulose-Murphy et al., 2001) and MHV-68 (Ahn et al., 1999) using microarrays.

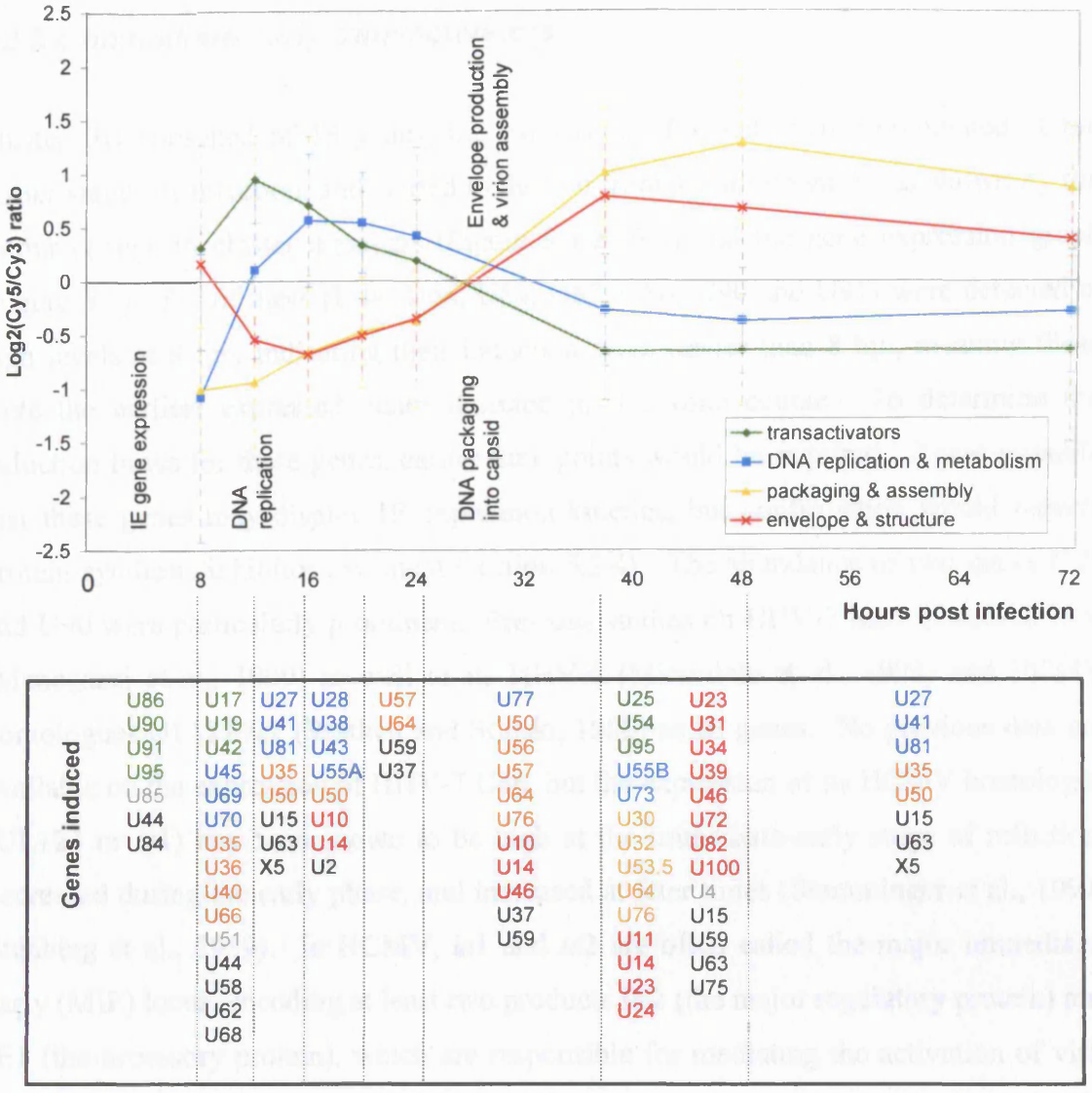


Figure 5.7 Mean expression ratio of HHV-7 genes by predicted functional group. The mean expression ratios of predicted transactivators (green), genes involved in DNA replication and metabolism (blue), in DNA packaging and capsid assembly (yellow), and envelope & structure (red) were taken from hierarchical clustering after mean-centring, and plotted over time. Error bars (dotted lines of corresponding colours) represent standard deviation. HHV-7 ORFs induced at various time points were colour coded and listed below the graph.

5.2.3.4 Immediate-early transactivators

Cluster IIB consisted of 15 genes, the expression of which were upregulated at the earlier stages of infection, and started to decline from 12 hpi onwards, as shown by the colour change in cluster diagrams (Figure 5.4 & 5.5) and the gene expression graph (Figure 5.7). Six of these (U44, U66, U84, U85, U86, U90 and U91) were detected at high levels at 8 hpi, indicating their induction were earlier than 8 hpi, meaning these were the earliest expressed genes detected in this time course. To determine the induction times for these genes, earlier time points would be required. It was possible that these genes may display IE expression kinetics, but confirmation would require protein synthesis inhibitor treatment (section 5.2.4). The abundance of two genes U86 and U90 were particularly prominent. Previous studies on HHV-7 have classified U90 (Menegazzi et al., 1999) as well as its HHV-6 (Mirandola et al., 1998) and HCMV homologues UL122/*ie*2 (Wathen and Stinski, 1982) as IE genes. No previous data are available on the expression of HHV-7 U86, but the expression of its HCMV homologue (UL123 or *ie*1) has been shown to be high at the immediate-early stage of infection, decreased during the early phase, and increased at later times (Stamminger et al., 1991; Stenberg et al., 1989). In HCMV, *ie*1 and *ie*2 are often called the major immediate-early (MIE) locus, encoding at least two products IE2 (the major regulatory protein) and IE1 (the accessory protein), which are responsible for mediating the activation of viral gene expression and regulating host gene expression (Spector, 1996; Stenberg, 1996). The MIE genes are expressed at relatively high levels compared to the genes of the other immediate-early (UL36-38) loci (Chambers et al., 1999), homologous to the IE-B (U17-U19) region in HHV-6 and HHV-7. Our results showed that this may also be true in HHV-7: U86 and U90 probes gave higher signals than U17 - U19 (Figure 5.3). Of the HHV-7 IE-B genes, U19 was assigned to cluster IIA (Figure 5.5a), whereas U18 was excluded from clustering analysis due to low signal. Interestingly, the expression profile of U17 was more similar to U86 and U90 than to U19, as it was assigned to cluster IIB in the 80%-presence cluster (Figure 5.5b).

Some transcription regulators were clustered elsewhere. For example, U95 was assigned to cluster I, showing a 'trough-shaped' expression profile, with high expression ratios at 8 hpi, which declined from 8 hpi to 48 hpi. This was followed by a

increase between 48 and 72 hpi (Figure 5.6a), possibly representing the beginning of a secondary infection. U42 was clustered with DNA replication genes, and U54 with the structural genes. These will be discussed in the relevant sections. Overall, the microarray-derived expression profile conformed to the known data on betaherpesvirus IE gene expression.

5.2.3.5 DNA replication and metabolism

The majority HHV-7 genes with predicted roles in DNA replication were found in clusters IIA and IIIA. These included DNA polymerase (U38), DNA polymerase processivity factor (U27), single-stranded DNA-binding protein (U41), and primase (U43). Helicase (U77), a second component of the heterotrimeric helicase-primase complex, was found in both clusters I and IIIB, whereas the third member, helicase-primase associated factor (U74), was not included in clustering. In addition, several genes with nucleotide metabolism and repair functions, hence indirectly involved in DNA replication, were also clustered with the DNA replication genes. These included ribonucleotide reductase (U28), dUTPase (U45), phosphotransferase (U69), alkaline exonuclease (U70), uracil-DNA glycosylase (U81) and U55A, a gene homologous to HCMV UL84, of which the product was believed to facilitate HCMV *ori*Lyt-dependent replication (Sarisky and Hayward, 1996). A gene expression transactivator (U42) was also clustered with the DNA replication-related genes. Its homologue in HCMV (UL69) may participate in DNA replication by upregulating the expression of DNA replication-related genes. The presence of HCMV UL69 product dramatically increases the transactivation of the TRS1 or IRS1 promoters (Romanowski and Shenk, 1997), which cooperate with $IE1_{491aa}$ and $IE2_{579aa}$ to transactivate E genes (Iskenderian et al., 1996; Romanowski and Shenk, 1997; Stasiak and Mocarski, 1992) of DNA replication functions (Pari et al., 1993). UL69 mutant virus growth was delayed under high MOI conditions due to an impact on DNA replication (Hayashi et al., 2000). We have found that HHV-7 U42 was highly expressed at the beginning of DNA replication stage (12 hpi), consistent with its probable function in transactivating enzymes and factors for viral DNA replication.

The expression of these DNA replication-related genes started low at 8hpi, then increased rapidly from 8 to 16 hpi (Figure 5.7), suggesting viral DNA replication

probably initiated during these 8 hours. These genes were subsequently downregulated between 24 and 38 hpi (Figure 5.7), suggesting that HHV-7 infection switched from a DNA replicating stage to a virion assembly stage between 24 and 38 hpi. Therefore expression profiling identified two major stages of HHV-7 infection cycle in human T-cells. After the reduction in expression between 24 and 38 hpi, some of these genes in cluster IIA (U19, U36, U42, U45 U69, U70) showed a second climb between 48 and 72 hpi (Figure 5.8b). This could be due to the beginning of a second round infection in new cells by the virions produced from the first round infection and may reflect asynchrony in the culture system after 48 hpi.

However, some DNA replication-related genes were assigned to other clusters. For example, only one each of the U43 and U70 array spots was assigned to cluster IIA, the other two were found in cluster IIB, a possible mis-assignment due to the missing values at 8 hpi, therefore the cluster IIA placement was probably more accurate. UL112/113, the HCMV homologue of HHV-7 U79, was suggested to have roles in DNA synthesis due to its ability to recruit replication-fork proteins to nuclear sites associated with viral transcription and DNA synthesis (Ahn et al., 1999; Penfold and Mocarski, 1997). However, hierarchical clustering assigned U79 to cluster I due to its 'trough-shaped' expression profile (Figure 5.6a). Although the functional significance of placement in cluster I was unclear, the relatively high abundance of U79 at 8hpi suggested that it was induced as early as the ORFs predicted as IE genes, and that it may have previously unknown functions in the immediate-early stages of infection. Another unexpected finding was that U55B had a very different transcription profile to U55A and was placed into cluster IIIB with the structural genes, despite the two genes were related (amino acid identity of 25%). The hybridisation signals on the microarrays were probably specific, as although the two ORFs were similar by amino acid sequence, their nucleotide sequences alignment matched no longer than 6 consecutive bases (Figure 5.9). Probes for U55A and U55B found no Blast match with the other ORF. Both HHV-7 U55A and U55B were suggested to have DNA replication-related functions, due to their homology to HCMV UL84, but our results showed that U55B may be involved in virion assembly stages, different to U55A. Perhaps the most surprising was that U73, believed to encode the origin binding protein, had an expression profile more similar to the structural genes in cluster IIIB than to the DNA

replication-related genes. However, the expression of U73 had two peaks (Figure 5.8a), suggesting second round of infection may have begun as early as 38 hpi.

5.2.3.6 Capsid assembly and Virion structure

Cluster IIIB was mainly made up of genes that were expressed at relatively low level during early stages of infection (8 – 24 hpi), followed by a major increase between 24 and 38 hpi (Figure 5.6e). Many of these genes are thought to encode structural proteins. These included genes involved in the formation of capsids, such as major capsid protein (U57), components of intercapsomeric triplex (U56, U76), and capsid protein located at the tip of hexons (U32); genes that encode virion envelope glycoproteins, gB (U39), gH (U82) and gM (U72); and other structural proteins such as glycoproteins (U23, U24, U100), large tegument protein (U31), virion proteins (U34, U50), antigenic tegument proteins (U11, U14), phosphoproteins (U10, U37), and a membrane protein (U46). In addition, genes involved in DNA packaging and capsid assembly, such as major capsid scaffold protein (U53.5), tegument proteins (U30, U64) were also assigned to this cluster. These findings suggested that HHV-7 DNA packaging and virion assembly initiated some time between 24 and 38 hpi during HHV-7-infection of SupT1 cells. The fact that genes encoding capsid, tegument and envelope proteins were clustered together indicated that proteins required for the formation of capsid, tegument and envelope were not expressed in three consecutive steps but simultaneously (Figure 5.7). The co-expression of all structural genes as one group has been observed in KSHV (Jenner et al., 2001).

Clustered with the structural genes was U54, predicted to encode a transactivator by homology to HCMV UL82. ppUL82 is an abundant tegument protein able to activate the *ie1*/*ie2* promoter-enhancer in transient transfection assays (Liu and Stinski, 1992), but also has activities independent of *ie1*/*ie2* expression (Baldick et al., 1997). UL82 expression starts early but peaks at late times in the infection (Gibson, 1983), consistent with our transcription profiling results on HHV-7 U54. HCMV UL82 is believed to act as a general transactivator of gene expression rather than as a specific activator of *ie1*/*ie2* gene expression (Mocarski, 2001). It is unknown what particular genes are

activated by HCMV UL82, but our data suggested that HHV-7 U54 may be a transactivator of structural genes.

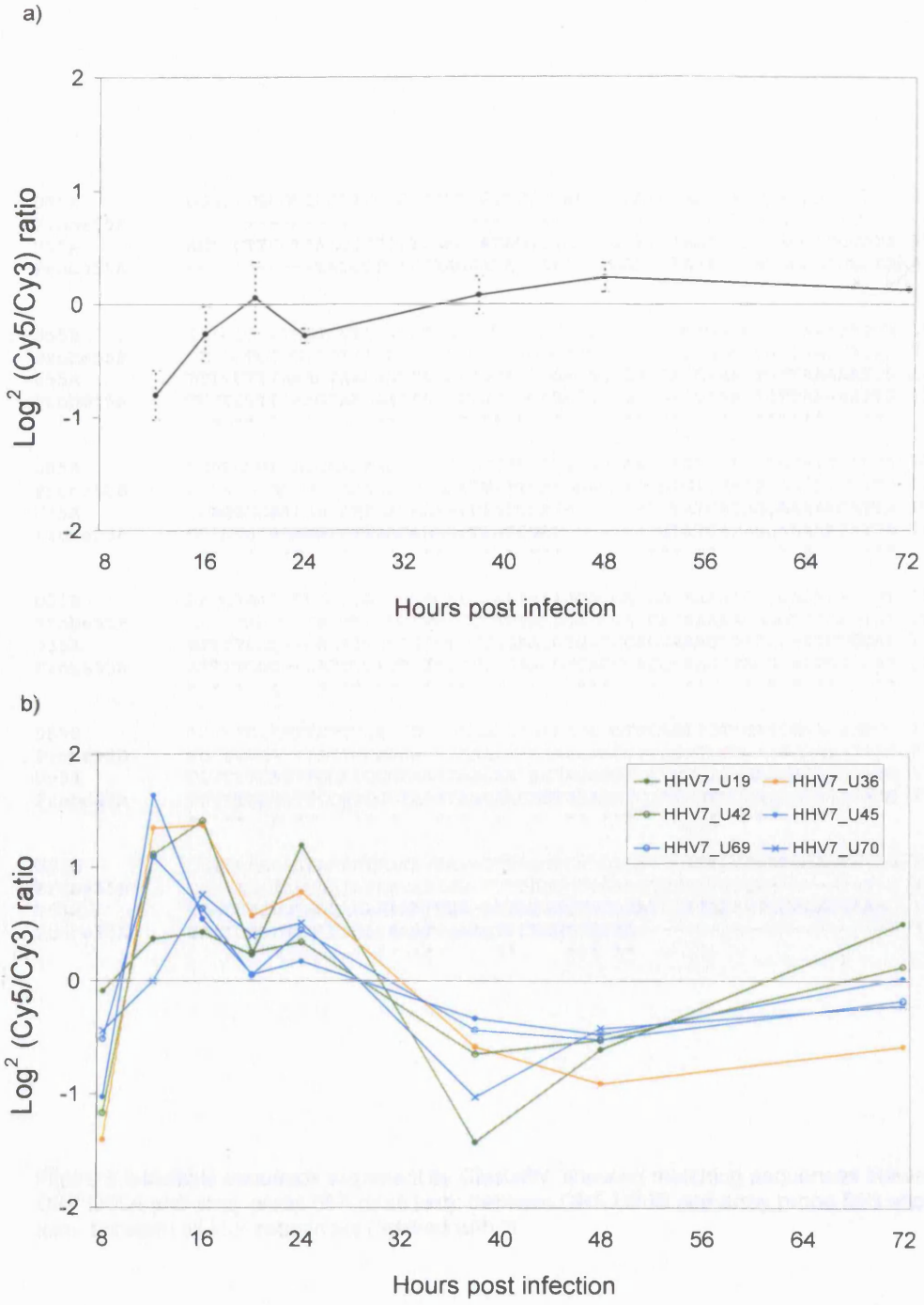


Figure 5.8 Mean expression ratio of a) HHV-7 U73 and b) selected cluster IIA genes during time course infection. Error bars (dotted lines) indicate standard deviation.

```
U55B        GGGTTTGCGTCACTATGACCTTTTCGTTTTTGAATCTTATCAGACTAGGAAAAACTCCCA 953
Probe55B    ------------------------------------------TGGAAAAACTCCCA 14
U55A        AGTTGTTCCTAAACCCTTTTCAAG-ATAGGCAAATTG-ATGAAATCTAAGAGGATCCATA 955
Probe55A    ----------TACCCTTTTCAAGGATAGGCAAATTGGATGAAATCTAAGAGGATCCATA 49
                                                  *  *  *  *

U55B        TCTGCTCTTGTTCTTTAACATCATGGTTTGGGATGATGACGATAGAGCTGTTAATTTTTA 1013
Probe55B    TCTGCTCTTGTTCTTTAACATCATGGTTTGGGATGATGACGATAGAGCTGTTAATTTTTA 74
U55A        TTTTCTTTAAAGTAACAATTATCTTGTATCAG-CAGGGATGACGTAA-TGTTAAAAATTG 1013
Probe55A    TTTTCTTTAAAGTAACAATTATCTTGTATCAGACAGGGATGACGTAA-TGTTAAAAATTG 108
            * * ** *        **     * ** *  *     ** **   *  ******   **

U55B        ACAACGGAAATCCAAAAAACAAGACTTTTGATGGATAAATGTGTCTATTCAGATTTTTTA 1073
Probe55B    ACAACGGAAATCCAAAAAACAAGACTTTTGATGGATAAATGTGTCTATTCAGATTTTTTA 134
U55A        GCAGGGCAAAAATATTGACAAAGTTTATCGAT-------TGTATCATAGGAAAAACATTA 1066
Probe55A    GCAGGGCAAAAATATTGACAAAGTTTATCGAT-------TGTATCATAGGAAAAACATTA 161
             **  * ***   *   *  ***  * * ***       *** **     * *    ***

U55B        AATGTAGTTTTAGTTGTTGTTCTTCCATTAAAGGACGAGCAGAAAAAAAGGAGTCAGAAT 1133
Probe55B    AATGTAGTTTTAGTTGTTGTTCTTCCATTAAAGGACGAGCAGAAAAAAAGGAGTCAGAAT 194
U55A        ATTTTGAT--CACTTTTTTGTTTTTTAGAACGTGACCCACCAAAGTTACGCATTCTGCAT 1124
Probe55A    ATTTTGAT--CACTTTTTTGTTTTTTAGAACGTGACCCACCAAAGTTACGCATTCTGCAT 219
            * * *  *   * ** **  * **  *  *   ***   *  **   * * * ** * **

U55B        TCCTTGAATTTATTTCAGCGCTCAAAAGTAAAACTGGTTCAGTTTTTGATTGGACATACT 1193
Probe55B    TCCTTGAATTTATTTCAGCGCTCAAAAGTAAAACTGGTTCAGTTTTTGATTGGACATACT 254
U55A        TGTTTTCAGTTCCATCGGTAATTAGCAACAGTACAGGTCTTTCGTGCGAAAGTATGCAGG 1184
Probe55A    TGTTTTCAGTTCCATCGGTAATTAGCAACAGTACAGGTCTTTCGTGCGAAAGTATGCAGG 279
            *  **  * **   ** *   * *  *  *  ** ***      *  **  * *   *

U55B        CGTCAAATAATAATGTAACATTAT-CTTTACCCTGGAATTCTTCTTTGTATGGAATTTCA 1252
Probe55B    CGTCAAATAATAATGTAACATTAT-CTTTACCCTGGAATTCTTCTTTGA---------- 302
U55A        TAATCTCTGCGGGAGAAAAATTGA-GTTCTGGCTGCAAATCATTAAAATCCACAGTAAAC 1243
Probe55A    TAATCTCTGCGGGAGAAAAATTGAAGTTCTGGCTGCAA---------------------- 317
                *     * ** ***    **    *** **
```

Figure 5.9 Multiple sequence alignment by ClustalW, showing matching sequences between HHV-7 ORF U55A and array probe 55A (blue text); between ORF U55B and array probe 55B shown as (red text); between all four sequences (marked with *).

5.2.3.8 Classification of HHV-7 genes from clustering pattern

The 68 HHV-7 genes were assigned to classes A, B, or C, based on clustering results (Table 5.3). Thirteen genes in clusters I and IIB, which showed a decreasing relative expression over time, were assigned to class A. However, probes with missing values at 8hpi were probably mis-assigned, and were therefore manually placed into Class B, which consisted of 24 genes that had low expression levels at 8 hpi, and were upregulated during 12 – 24 hpi (clusters IIA and IIIA). Finally, 25 genes that were induced later in the time course (cluster IIIB) were placed into class C. Ten genes displayed expression profiles that could belong to more than one class, thus the classification of these await further analysis using metabolic inhibitors (section 5.2.4 – 5.2.5). Genes in classes A, B and C may be equivalent to IE, E and L genes, respectively, but confirmation of this would require the use of metabolic inhibitors to block *de novo* protein synthesis and DNA replication (section 5.2.4 – 5.2.5).

Our classification of HHV-7 genes was compared to previous temporal mapping studies of HHV-7, HHV-6 and HCMV genes (Table 5.3). Two of the class A genes in this study (U17, U90) have been previously analysed by RT-PCR with metabolic inhibitors and were reported to be IE-genes in unspliced form in HHV-7 (Menegazzi et al., 1999) and in HHV-6B (Mirandola et al., 1998). Of the class B genes, however, only one (U19) has been classified as and E-gene in both HHV-7 and HHV-6 (Menegazzi et al., 1999; Mirandola et al., 1998). U41 has been shown to be an IE-gene in HHV-7 (Menegazzi et al., 1999) and an E gene in HHV-6 (Mirandola et al., 1998); while U42 was classified as an IE-gene in both studies. U91 has been given both IE- and E-kinetics in unspliced and spliced form, respectively, in HHV-6A (Mirandola et al., 1998). Four genes (U14, U31, U39, U73) classified as class C in this thesis were shown to be IE-genes by Menegazzi *et al.* Our expression profiling results were also compared to a study of HCMV that used microarrays to detect transcripts in the presence of metabolic inhibitors (Chambers et al., 1999). Results were slightly more comparable, but not without discrepancies. Of the 24 class B genes in this study, the HCMV homologues of only six genes (U17, U38, U40, U41, U51, U81) have been classified as E-genes, while eight (U27, U42, U45, U55A, U66, U68, U69, U70) and two (U36, U62) were thought to be E-L and L genes, respectively (E-L genes were defined as late genes

with reduced expression, and L genes with undetectable expression in the absence of viral DNA replication). The class C genes were more consistent with the HCMV results, with the HCMV homologues of six (U11, U25, U31, U54, U64, U75) and four (U30, U56, U57, U72) HHV-7 genes also classified as L- and E-L genes, respectively. Nonetheless, five class C genes (U4, U14, U37, U39, U76) had E-kinetics in HCMV. Chambers *et al* only identified one IE gene (UL122) that has a counterpart in HHV-7 (U86), which was found in class A in our study. Another class A gene (U79) however, was considered to be an E/E-L gene in HCMV (Chambers et al., 1999).

These findings suggested that data obtained from microarray experiments are more comparable to each other than to RT-PCR studies. However, the reason for this may be simply because microarray studies generate more data to be compared. Our assignment of HHV-7 genes into classes A, B and C were based on the transcription profile of virus-infected cells at a high MOI of 2.5, in a relatively synchronous, natural infection *in vitro*; whereas treatment with metabolic inhibitors have damaging effects on the cells that may influence viral transcription (Honess and Watson, 1977).

5.2.3.9 Predicting functions for previously uncharacterised genes

Ten (U37, U44, U47, U58, U59, U62, U63, U68, U75, U84) of the HHV-7 genes analysed by clustering had no known function. The roles of these 10 genes could be predicted from the time at which they were induced (Figure 5.7), and from the clustering pattern (Figure 5.4, 5.5), because genes with similar functional roles were often grouped together by clustering analysis. U84 was induced before 8hpi and had similar expression pattern to the class A transactivators, therefore may itself be a transactivator involved in the immediate-early stages of infection. Several genes (U2, U44, U62, U68, U63) may have roles in viral DNA replication, as they were expressed at highest levels during 12 – 24 hpi and were clustered with DNA replication-related genes. Four ORFs (U59, U37, U63, U75) may encode for structural products or proteins involved in DNA packaging, inferred from the late induction (after 24 hpi) of their gene expression, which was typical amongst structural genes.

5.2.3.10 Summary of transcription profiling

In summary, monitoring HHV-7 gene expression by microarrays followed by transcription profiling using hierarchical clustering has revealed the regulated pattern of HHV-7 gene expression during infection of SupT1 cells. It has also allowed the classification of HHV-7 genes into classes A, B and C, based on their expression patterns, which also represented three stages of viral infection cycle. Figure 5.7 showed the proposed HHV-7 infectious cycle in SupT1 cells *in vitro*. After virus attachment and internalisation, expression of class A genes began within 8 hours, which include many genes believed to act as IE transactivators. Most genes involved in DNA replication and metabolism were found in class B. These were mostly induced between 8 and 16 hpi, and their expression started to decline after 24 hpi, to be taken over by the transcription of class C genes, which were involved in DNA packaging, capsid assembly and virion formation. The second round of DNA replication began between 48 and 72 hpi, suggesting the beginning of second round infection with the virions released from first round infection took place as early as 36 hpi (= 48 – 12) and by 60 hpi (= 72 – 12). Our results therefore suggested that the duration of infectious cycle of HHV-7 in SupT1 is between 36 to 60 hours, shorter than the 4-day cycle in cord blood lymphocytes previously suggested (Black et al., 1997). Analysis of infected cells at a time point between 48 and 72 hpi (e.g. 60 hpi) may facilitate the identification of a more precise duration.

Genes in classes A, B and C also roughly corresponded to genes previously classified as IE, E and L genes, respectively, in betaherpesviruses. Therefore expression profiling provided insight into the co-regulated expression of viral genes during different stages of its infection cycle in T-cells, and was able to classify genes into their temporal classes with results comparable with published findings.

5.2.4 Temporal classification of HHV-7 genes using metabolic inhibitors

To classify HHV-7 genes into kinetic classes of immediate early (IE), early (E) and late (L) genes, SupT1 cells were treated with metabolic inhibitors for 1 hour before infection by HHV-7. Cells treated with 200 μg/ml cycloheximide (CHX) to block *de novo* protein synthesis were harvested at 8 hpi; cells treated with 500 μg/ml phosphonoacetic acid (PAA) to block viral DNA replication were harvested at 24 and 48 hpi. The expression profiles of inhibitor-treated cells were analysed by the custom-made host-pathogen microarrays followed by clustering analysis (80% presence) in the same way as the untreated cells (section 5.2.3).

5.2.4.1 Immediate-early and Early genes

At 8 hpi, treating with CHX probably had little effect on viral gene expression patterns, as the expression profiles for both CHX-treated and untreated cells at 8hpi were highly similar and clustered together (Figure 5.10a). CHX was added to block protein synthesis in order to differentiate the IE genes that were independent of protein synthesis from the E and L genes that required *de novo* protein synthesis for their transcription. There could be two possible reasons for the lack of differentiation between CHX-treated and untreated cells at 8 hpi. First, the CHX might be ineffective in blocking protein synthesis in the HHV-7 infected SupT1 cells. Second, the expression of protein synthesis-dependent genes had not started or was not detectable by 8 hpi. The first reason was unlikely, because the presence of CHX indeed inhibited the expression of many proposed E and L genes (section 5.2.5). The second reason was probable, as it had been described in section 5.2.1 (Figure 5.4 – 5.5) that the majority of viral early transcripts only became detectable some time between 8 and 12 hpi. In order to prove this and to reveal the distinction between expression patterns in CHX-treated and untreated cells, a later time point for CHX-treated cells would be required. A 16 hpi sample has been attempted, but most cells have been killed in the presence of CHX by this time, therefore it was not possible to extract sufficient RNA for array analysis. In future experiments, a CHX-treated sample should be harvested at 12 hpi, to

determine if more genes differentially expressed between treated and untreated cells could be detected. Moreover, at 8hpi the abundances of many genes were below the signal-to-noise filtering threshold thus were excluded from analysis. Therefore we were unable to differentiate between IE and E genes using our microarrays. The lack of sensitivity for identifying IE genes by microarrays has been noticed in previous microarray study of HCMV (Chambers et al., 1999).

5.2.4.2 Early and Late genes

Since there were little difference between CHX-treated and untreated cells at 8 hpi, also 8 hpi samples contained a lot of missing data, clustering analysis was next performed without the 8 hpi samples, and focussed on the later samples, in order to optimise the number of genes analysed (Figure 5.10b). HHV-7 genes were classified into IE/E or L genes according to their response to PAA treatment. The expression levels of 15 genes were inhibited by PAA treatment (Figure 5.10b, 5.12), hence they were classified as viral DNA replication-dependent, or L genes. The differential expression of 8 of these L genes (U23, U46, U53.5, U54, U55B, U56, U57, U82) between treated and untreated cells were statistically significant, as determined by Significance analysis of microarrays (SAM) with false discovery rate set at 3.0%, or 0.54 gene (Figure 5.11). All remaining genes were classified as IE/E genes (Table 5.3), as their expression ratios were not reduced in the presence of PAA hence their transcription were independent of viral DNA replication (Figure 5.12).

Figure 5.10 (next page) a) Hierarchical clustering analysis of HHV-7 gene expression at 8, 24, 48 hours post infection of SupT1 cells in the presence and absence of metabolic inhibitors that block de novo protein synthesis (cycloheximide, CHX); and viral DNA replication (phosphonoacetic acid PAA). b) Hierarchical clustering analysis excluding 8 hpi samples; right panel, zoomed image of the cluster containing genes downregulated in the presence of PAA.

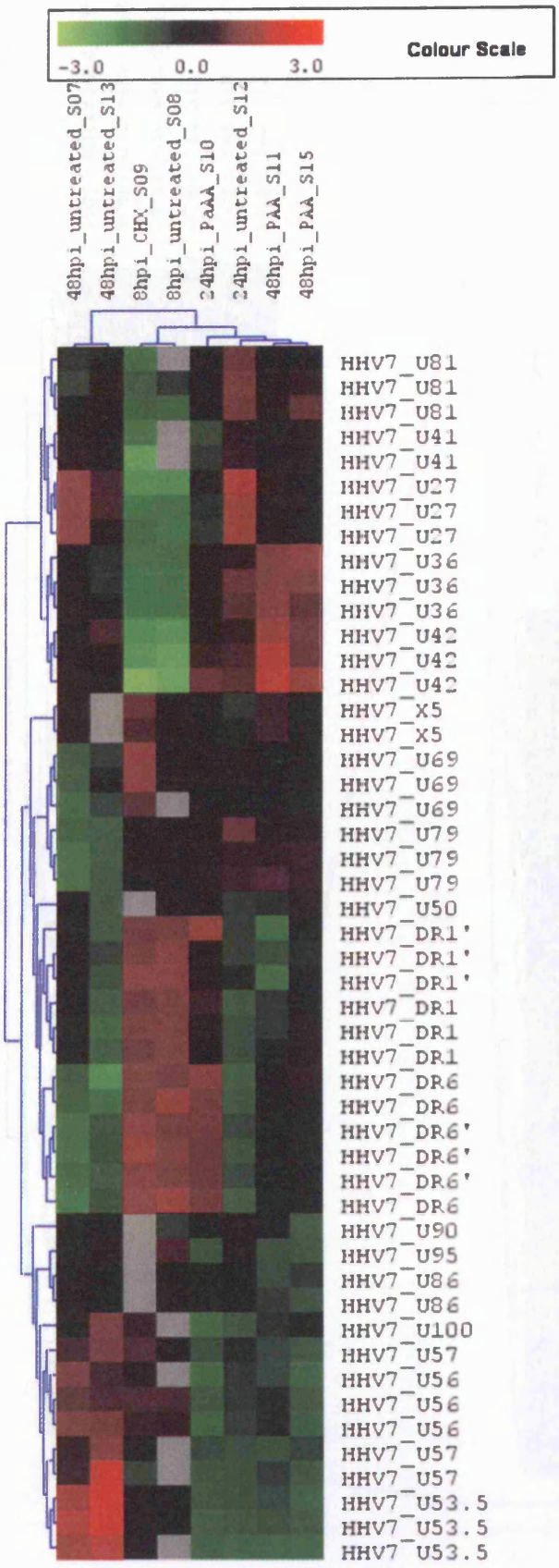

Colour Scale
-3.0
0.0
3.0
48hpi_untreated_S07
48hpi_untreated_S13
8hpi_CHX_S09
8hpi_untreated_S08
24hpi_PAA_S10
24hpi_untreated_S12
48hpi_PAA_S11
48hpi_PAA_S15
HHV7_U81
HHV7_U81
HHV7_U81
HHV7_U41
HHV7_U41
HHV7_U27
HHV7_U27
HHV7_U27
HHV7_U36
HHV7_U36
HHV7_U36
HHV7_U42
HHV7_U42
HHV7_U42
HHV7_X5
HHV7_X5
HHV7_U69
HHV7_U69
HHV7_U69
HHV7_U79
HHV7_U79
HHV7_U79
HHV7_U50
HHV7_DR1'
HHV7_DR1'
HHV7_DR1'
HHV7_DR1
HHV7_DR1
HHV7_DR1
HHV7_DR6
HHV7_DR6
HHV7_DR6'
HHV7_DR6'
HHV7_DR6'
HHV7_DR6
HHV7_U90
HHV7_U95
HHV7_U86
HHV7_U86
HHV7_U100
HHV7_U57
HHV7_U56
HHV7_U56
HHV7_U56
HHV7_U57
HHV7_U57
HHV7_U53.5
HHV7_U53.5
HHV7_U53.5

Figure 5.10a

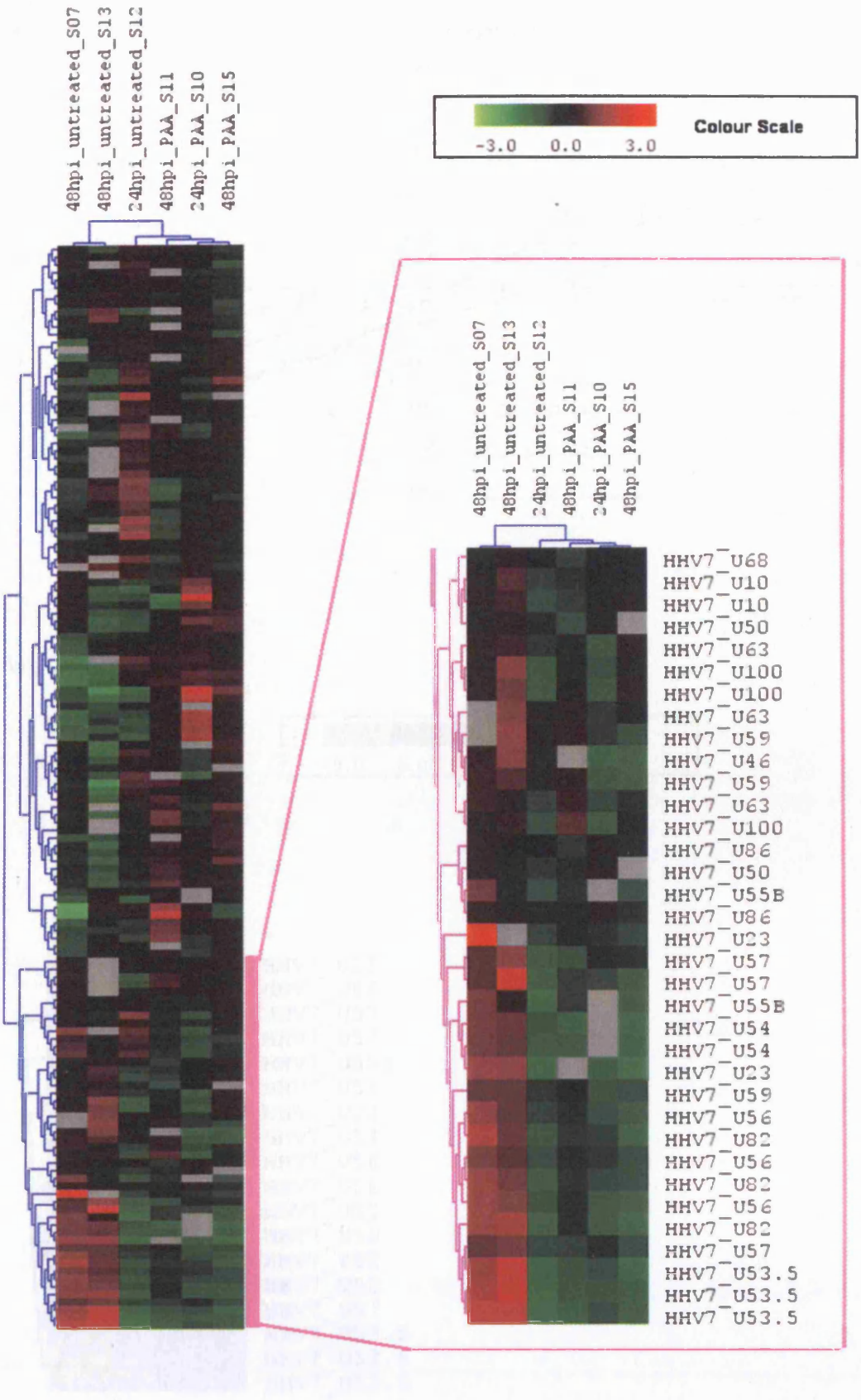

48hpi_untreated_S07
48hpi_untreated_S13
24hpi_untreated_S12
48hpi_PAA_S11
24hpi_PAA_S10
48hpi_PAA_S15
Colour Scale
-3.0
0.0
3.0
HHV7_U68
HHV7_U10
HHV7_U10
HHV7_U50
HHV7_U63
HHV7_U100
HHV7_U100
HHV7_U63
HHV7_U59
HHV7_U46
HHV7_U59
HHV7_U63
HHV7_U100
HHV7_U86
HHV7_U50
HHV7_U55B
HHV7_U86
HHV7_U23
HHV7_U57
HHV7_U57
HHV7_U55B
HHV7_U54
HHV7_U54
HHV7_U23
HHV7_U59
HHV7_U56
HHV7_U82
HHV7_U56
HHV7_U82
HHV7_U56
HHV7_U82
HHV7_U57
HHV7_U53.5
HHV7_U53.5
HHV7_U53.5

Figure 5.10b

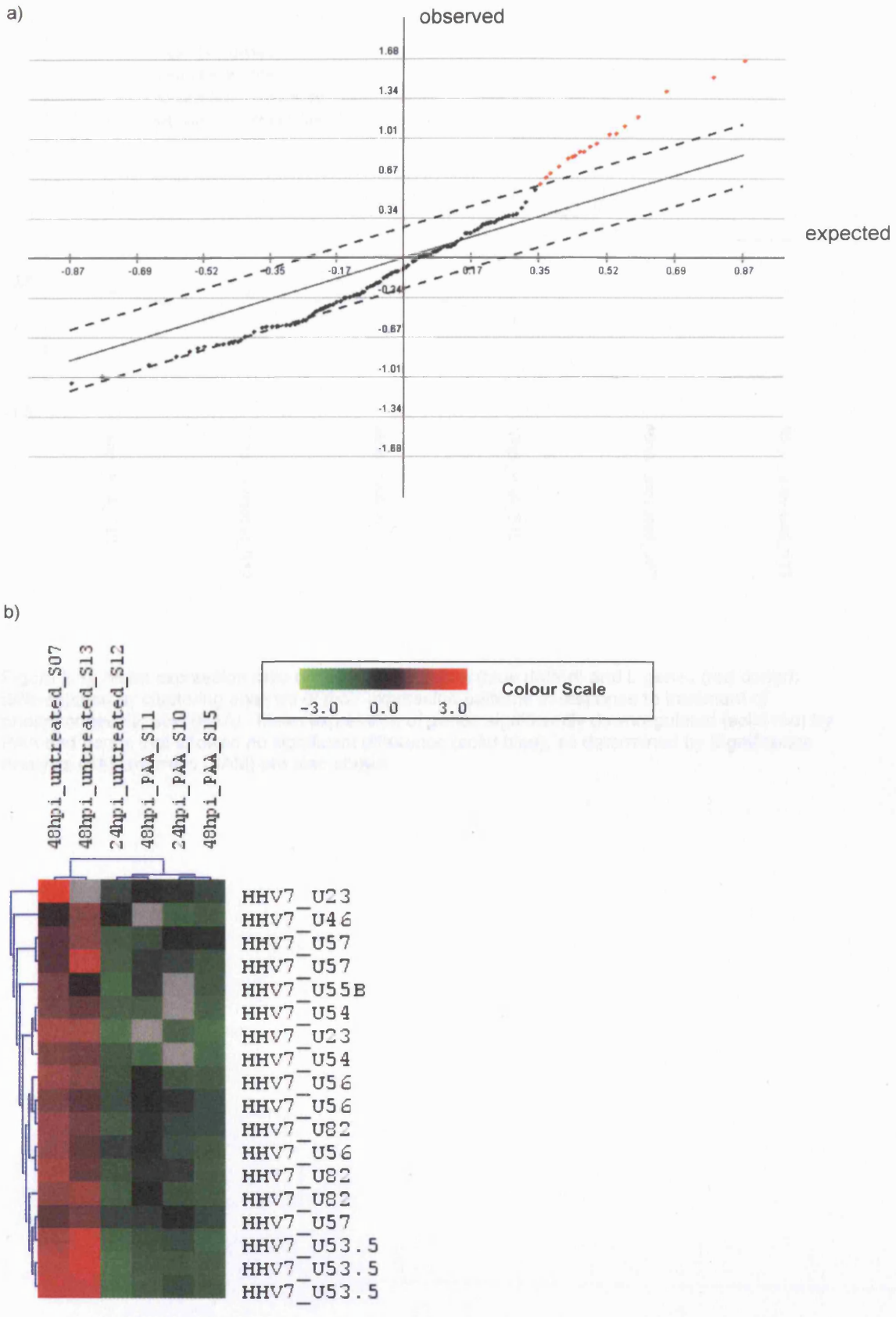


Figure 5.11 Significance analysis of microarrays (SAM) was used to identify statistically significant downregulated HHV-7 genes in the presence of phosphonoacetic acid (PAA) during infection of SupT1 cells. a) SAM scatter plot of observed (y-axis) vs expected (x-axis) expression values of genes significantly downregulated by PAA treatment (red dots) and unchanged genes (black dots). b) Genes significantly downregulated by PAA were clustered by hierarchical clustering.

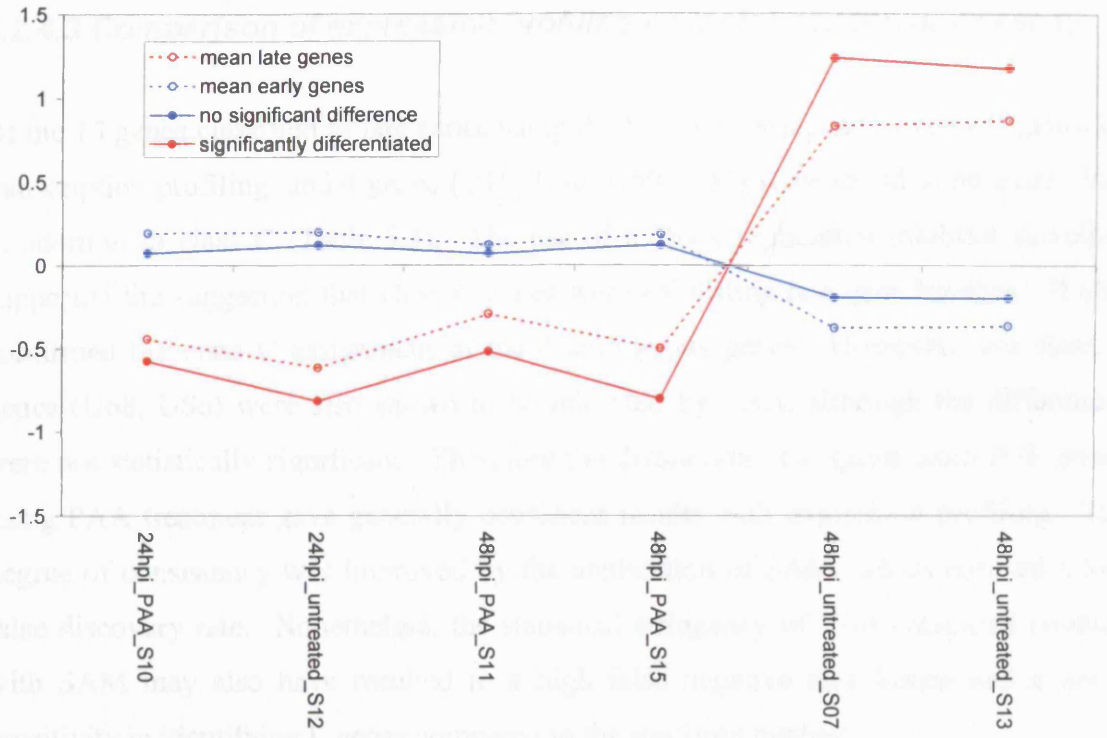


Figure 5.12 Mean expression ratio of HHV-7 IE/E genes (blue dotted) and L genes (red dotted) differentiated by clustering analysis of their expression patterns in response to treatment of phosphonoacetic acid (PAA). Mean expression of genes significantly downregulated (solid red) by PAA and genes that showed no significant difference (solid blue), as determined by Significance Analysis of Microarrays (SAM) are also shown.

5.2.4.3 Comparison of expression profiling vs metabolic inhibitor assay

Of the 15 genes classified as late genes using PAA, 9 were assigned as class C genes by transcription profiling, and 4 genes (U10, U46, U50, U63) were found in an extra class in addition to class C (Table 5.3). The use of a DNA replication inhibitor therefore supported the suggestion that class C genes were exhibiting late gene kinetics. It also confirmed the class C assignment of the 4 ambiguous genes. However, two class A genes (U68, U86) were also shown to be inhibited by PAA, although the differences were not statistically significant. Therefore the distinction of L genes from IE/E genes using PAA treatment gave generally consistent results with expression profiling. The degree of consistency was improved by the application of SAM, which ensured a low false discovery rate. Nonetheless, the statistical stringency of PAA-treatment coupled with SAM may also have resulted in a high false negative rate, hence had a lower sensitivity in identifying L genes compared to the profiling method.

ORF	Predicted function	This thesis		Published results				HCMV homologue ORF [f]	HCMV Array [g]
		transcription profiling [a]	metabolic inhibitors [b]	HHV-7 RT-PCR [c]	HHV-6A RT-PCR [d]	HHV-6B RT-PCR [d]	HHV-6B Array [e]		
DR1	US22 gene family	A	IE/E				L		
DR6	US22 gene family	A	IE/E				L		
U2	US22 gene family	B					L	UL23	
U3	US22 gene family							UL24	
U4	related to U7 exon 2	C					L	UL27	E
U7	US22 gene family, related to U4						L	UL27	E
U8	US22 gene family		IE/E				L	UL29	L
U10	nuclear phosphoprotein	C	L	IE				UL31	L
U11	Structural phosphoprotein	C						UL32	L
U12	G-protein coupled receptor, chemokine receptor		IE/E					UL33	E
U13									
U14	HCMV UL25/35 gene family, tegument phosphoprotein	C		IE			L	UL35	E
U15		B/C	IE/E						
U17	IE-B, transactivator, US22 gene family	B	IE/E	IE (unspliced); E (spliced)	E	IE		UL36x1	E
U18	IE-B, homologue to HCMV IE protein			IE	E	E		UL37x3	
U19	IE-B protein	B	IE/E	E	E	E	L	UL38	
U20	probable membrane glycoprotein, Ig gene family?			E	E	E			
U21	probable membrane glycoprotein		IE/E						
U23	probable membrane glycoprotein, EHV-1 gJ homologue	C	L						
U24	Glycoprotein	C							
U24A	contains hydrophobic domain								
U25	US22 gene family, transactivator	C					L	UL43	L
U26							L		
U27	DNA polymerase processivity factor	B						UL44	E-L
U28	Ribonucleotide reductase (large subunit)	B	IE/E					UL45	
U29	Minor capsid protein (mCP)							UL46	E-L
U30	Capsid assembly, myosin	C						UL47	E-L
U31	Large tegument protein	C		IE	E	E	L	UL48	L
U32	Capsid protein (located on tips of hexons)	C					L	UL48.5	
U33	Virion protein		IE/E					UL49	E-L

ORF	Predicted function	This thesis		Published results				HCMV homologue ORF [f]	HCMV Array [g]
		transcription profiling [a]	metabolic inhibitors [b]	HHV-7 RT-PCR [c]	HHV-6A RT-PCR [d]	HHV-6B RT-PCR [d]	HHV-6B Array [e]		
U34	Virion protein?	C						UL50	
U35	Role in DNA packaging	B	IE/E				L	UL51	
U36	Probable virion protein	B	IE/E				L	UL52	L
U37		C						UL53	E
U38	DNA polymerase	B	IE/E				L	UL54	E
U39	Envelope glycoprotein B (gB)	C	IE/E	IE	IE	IE	L	UL55	E
U40	Transport protein TP	B	IE/E					UL56	E
U41	Major DNA-binding protein,	B	IE/E	IE	E	E	L	UL57	E
U42	Transactivator	B	IE/E	IE	IE	IE	L	UL69	E-L
U43	Component of DNA helicase-primase complex, primase	B	IE/E				L	UL70	
U44		A	IE/E					UL71	
U45	dUTPase	B	IE/E				L	UL72	E-L
U46	Membrane protein	B/C	L				L	UL73	E-L
U47		A	IE/E					UL74	
U48	Envelope glycoprotein H (gH)						L	UL75	E-L
U49	Fusion protein							UL76	
U50	Virion protein	B/C	L					UL77	E
U51	G-protein coupled receptor	B	IE/E				L	UL78	E
U52	opioidR homologue						L	UL79	
U53	Protease/assembly protein			IE	E	E	L	UL80	
U53.5	Major capsid scaffold protein	C	L						
U54	Tegument protein transactivator	C	L				L	UL82/83	L
U55A	Replication function?	B	IE/E				L	UL84	E-L
U55B	Replication function?	C	L				L	UL84	
U56	Capsid protein	C	L				L	UL85	E-L
U57	Major capsid protein (MCP)	C	L				L	UL86	E-L
U58		B	IE/E				L	UL87	
U59		C	L				L	UL88	
U62		B						UL91	L
U63		B/C	L				L	UL92	L
U64	Role in DNA packaging, tegument protein	C	IE/E				L	UL93	L

ORF	Predicted function	This thesis		Published results				HCMV homologue ORF [f]	HCMV Array [g]
		transcription profiling [a]	metabolic inhibitors [b]	HHV-7 RT-PCR [c]	HHV-6A RT-PCR [d]	HHV-6B RT-PCR [d]	HHV-6B Array [e]		
U65	Tegument protein						L	UL94	L
U66	Late spliced gene (U60/U66) DNA packaging	B	IE/E	IE/L			L	UL89x2	E-L
U67			IE/E					UL95	
U68		B	L				L	UL96	E-L
U69	Phosphotransferase	B	IE/E				L	UL97	E-L
U70	Alkaline exonuclease	B	IE/E				L	UL98	E-L
U71	Myristylated tegument protein								
U72	Integral membrane protein (gM)	C					L	UL100	E-L
U73	Origin-binding protein (OBP)	C		IE	IE	IE	L		
U74	Component of DNA helicase/primase complex						L	UL102	L
U75		C					L	UL103	L
U76	Virion protein?	C						UL104	E
U77	Component of DNA helicase/primase complex	C	IE/E				L	UL105	E
U79	HCMV replicaton, spliced (UL112/UL113)	A	IE/E				IE/E	UL112/113	E/E-L
U81	Uracil-DNA glycosylase	B	IE/E					UL114	E
U82	Envelope glycoprotein L (gL)	C	L				IE/E	UL115	
U84	Spliced in HCMV	A	IE/E					UL117	
U85	probable membrane glycoprotein, related to OX-2	A							
U86	IE-A protein, HCMV IE2 homology	A					IE/E	UL122	IE, L
U90	IE-A transactivator	A	IE/E	IE	IE	IE	IE/E		
U91	probable membrane glycoprotein	A	IE/E		IE (unspliced); E (spliced)	E (unspliced); IE (spliced)			
U95	US22 gene family,	A	IE/E				IE/E		
U100	Homology to HHV-6 gp82/105	C	L	E/L	L	L	IE/E		
DR1'	US22 gene family	A							
DR6'	US22 gene family	A							
X1									
X3									
X5		B							
X7		A	IE/E						

Table 5.3 Classification of HHV-7 genes in this thesis and previous studies

[a] Transcription profiling of HHV-7 genes based on clustering pattern of time course samples (8,12,16,20,24,38,48,72 hpi); class A, genes in clusters I and IIB; class B, genes in cluster IIA and IIIA; class C, genes in cluster IIIB

[b] Statistical analysis of infected in the presense or absense of metabolic inhibitors: cyclohexide (8 hpi) and phosphonoacetic acid (24,48 hpi)

[c] Temporal classification of HHV-7 genes using RT-PCR and metabolic inhibitors (Menegazzi et al., 1999)

[d] Temporal classification of HHV-6A and HHV-6B genes using RT-PCR and metabolic inhibitors (Mirandola et al., 1998)

[e] Temporal classification of HHV-6B genes using arrays and metabolic inhibitors (Ohyashiki et a., 2005)

[f] HCMV homologues of HHV-7 (Nicholas, 1996; Megaw et al., 1998)

[g] Temporal classification of HCMV genes using microarrays and metabolic inhibitors (Chambers et al., 1999)

5.2.5 Temporal classification of HHV-7 genes by RT-PCR

5.2.5.1 RT-PCR conditions

To verify microarray results, reverse transcription polymerase chain reaction (RT-PCR) was performed for a selection of ORFs. RT-PCR reactions were carried out using the OneStep kit (Qiagen, UK), which contained a mixture of reverse transcriptase and heat-activated HotStar DNA polymerase (Qiagen, UK). Firstly, RNA was reverse-transcribed into cDNA at 50 °C for 30 min, followed by heating at 95 °C for 15 min to inactivate the reverse transcriptase. The heating also activated the DNA polymerase for 35 cycles of PCR amplification reaction of the cDNA. However, it was found that the reverse transcriptase was not completely inactivated by heating, because if RNA was added after the heating step, PCR products were still formed (Figure 5.13). In contrast, using HotStar DNA polymerase (Qiagen, UK) alone produced no products with RNA samples. The HotStar DNA polymerase was capable of amplifying a product from viral DNA (Figure 5.13), but not from total RNA or amplified RNA (aRNA) in the absence of reverse transcription, indicating that there was no DNA contamination in the RNA samples. Some Taq-only reactions yielded products for some ORFs (Figure 5.14), although some of these bands (U17, U90) corresponded to spliced mRNAs therefore they were probably cDNAs from the 1st strand synthesis step in the RNA amplification reaction, rather than DNA contamination.

Initially 50 ng of aRNA was used as template, which resulted in strong product bands but failed to differentiate the expression of a predicted L gene (U56) between untreated or PAA-treated cells at 48 hpi; or the expression of a predicted IE gene (U17) between untreated and CHX-treated cells at 8 hpi (Figure 5.13). This could be either because PAA and CHX treatments made no difference to the expression of these genes, or because at this high template concentration, the transcripts were amplified to saturation, masking any differences. The latter was possible because if the inhibitors were not 100% efficient, transcripts that were supposedly inhibited by PAA or CHX may be present at a reduced level, as it has been observed in the array results (section 5.2.4). The aRNA has already been amplified approximately 1000-

fold by *in vitro* transcription (IVT) (section 4.2.5), so the RT-PCR was effectively the second round of amplification. Unlike IVT, PCR amplifies exponentially and as a result, the transcripts may have been amplified to saturation, masking the original differences in abundance. In an attempt to prove and rectify this, RT-PCR was performed using a serial dilution of template (from 50 ng to 0.5 pg for U56 in 48hpi samples; or 50 ng to 1 ng for U17 in 8hpi samples), to check if a difference could be observed between inhibitor-treated and untreated samples at lower template concentrations (Figure 5.13). The optimum template concentration for differentiating U56 transcript abundance between PAA-treated and untreated cells at 48hpi was found to be 500 pg/ul. This demonstrated that U56 was indeed an L gene, whose expression was reduced by PAA. For U17 expression at 8hpi, sensitivity to CHX was observed at the template concentration of 10 ng/ul. A template concentration of 1 ng/ul was estimated for 24 hpi samples. These template concentrations were used in subsequent RT-PCR reactions for 12 HHV-7 ORFs.

5.2.5.2 Classification of HHV-7 gene expression by RT-PCR

Expression of the U17 gene was not affected by PAA thus it was not a late gene. CHX treatment reduced the intensity of the two smaller bands and caused the largest band to disappear. The bands were gel purified and sequenced. It was found that the middle band corresponded to genomic sequence hence was an unspliced transcript, whereas the smallest band was a spliced transcript (Appendix C). RT-PCR results thus showed that both spliced and unspliced versions of U17 were reduced by CHX treatment, thus U17 was considered an E gene in both forms. This was slightly different to the previous RT-PCR study, which suggested U17 produced unspliced IE and spliced E transcripts (Menegazzi et al., 1999). The large U17 band produced mixed species when sequenced. U90 expression was not affected by CHX or PAA thus was most likely expressed as an IE gene. The transcription product was smaller than the original probe, which was revealed to be a spliced product by sequence verification (Appendix C). This agreed with previous data (Menegazzi et al., 1999), although Menegazzi *et al* also detected an unspliced form of U90 as an IE transcript.

RT-PCR results showed U27, U39 and U41 to be IE genes, with expressions only slightly reduced by CHX and unaffected by PAA. This was consistent with previous RT-PCR study (Table 5.4) but not with expression profiling using microarray in this study (Table 5.3), which classified U27 and U41 to be class B genes, and U39 to be a class C gene, even though low signals were detected at 8hpi. Both U27 and U41 are thought to be involved in DNA replication. U27 is the DNA polymerase processivity factor and U41 the single-stranded DNA binding protein. From their functions one would predict E kinetics for their expression. However, our data suggested that some low level expression occurred during the IE stage of infection without the need for transactivators, but their expressions were highly accelerated after the IE products had accumulated. The alternative explanation was that CHX was not 100% effective in blocking *de novo* protein synthesis, therefore a low level 'leaky' expression took place.

U66 transcript was most abundant at 24hpi, with or without PAA treatment. It was also detected in untreated cells at 8hpi, but was inhibited by CHX. By 48 hpi the abundance had decreased, therefore U66 displayed a classic E gene profile. This was consistent with microarray classification (Table 5.3) but not with previous results using RT-PCR (Table 5.4). U38, which encodes the viral DNA polymerase, also displayed patterns indicative of E gene (Figure 5.14), consistent with microarray results but no previous RT-PCR data is available for comparison.

Expression of ORFs U10, U50, U53.3 U56 and U82 were all reduced by PAA treatment, suggesting L-kinetics (Figure 5.14). This was consistent with microarray data (Table 5.3). Where replicates of the genes U10 and U50 were present in more than one cluster, RT-PCR results helped towards the final decision of their classification as L genes.

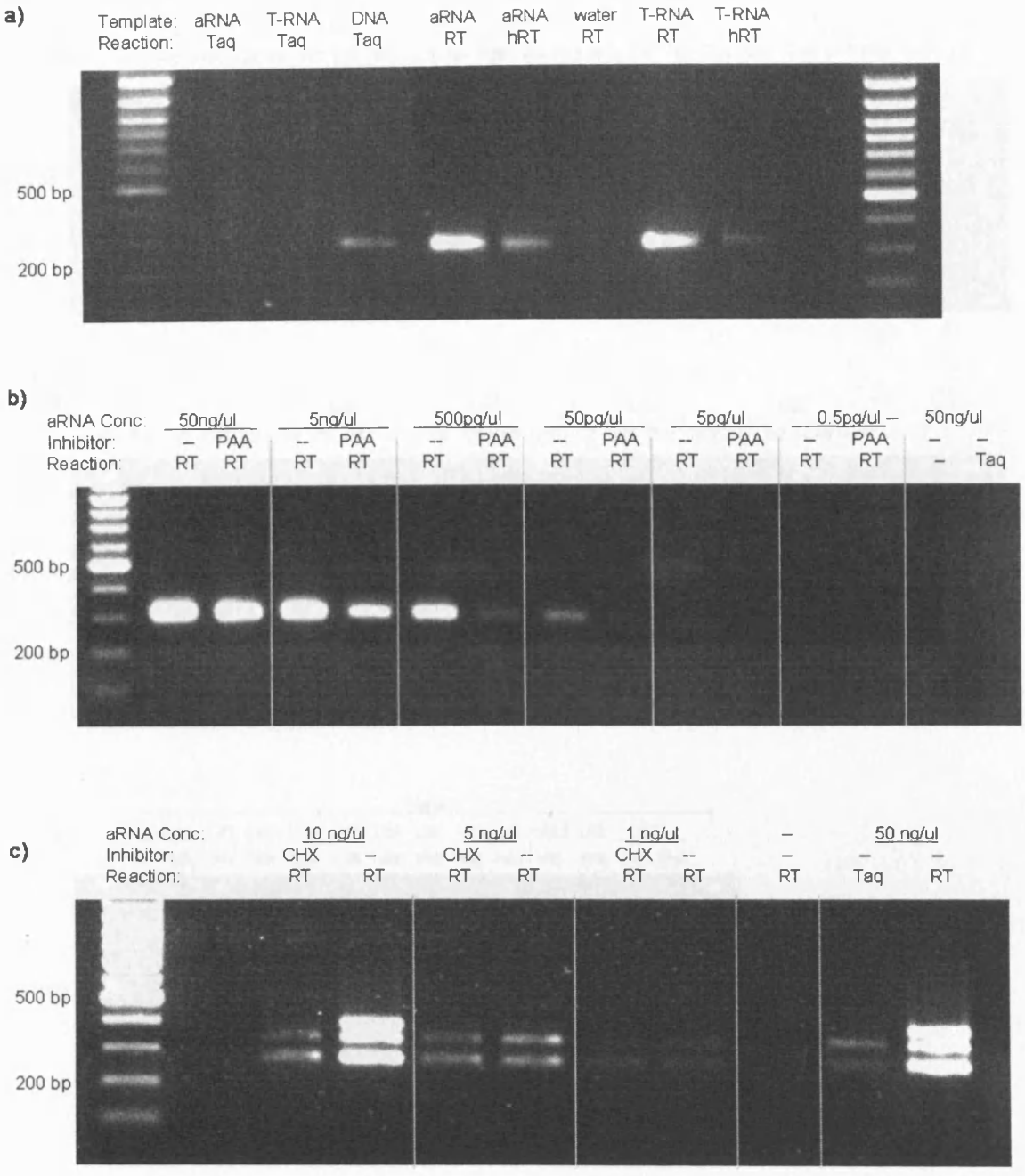


Figure 5.13 RT-PCR using specific primers for HHV-7 ORF a) U82, b) U56, c) U17; template: HHV-7 DNA as positive control; water as negative control; amplified RNA (aRNA) or total RNA (T-RNA) purified from HHV-7 infected SupT1 cells at a) 48hpi, b) 48hpi, c) 8hpi. Reactions were performed with either reverse transcriptase (RT), heat-inactivated reverse transcriptase (hRT), or Taq DNA polymerase (Taq). b-c) template aRNA was serial diluted, and were purified from cells treated with cycloheximide (CHX); phosphonoacetic acid (PAA); or untreated (--).

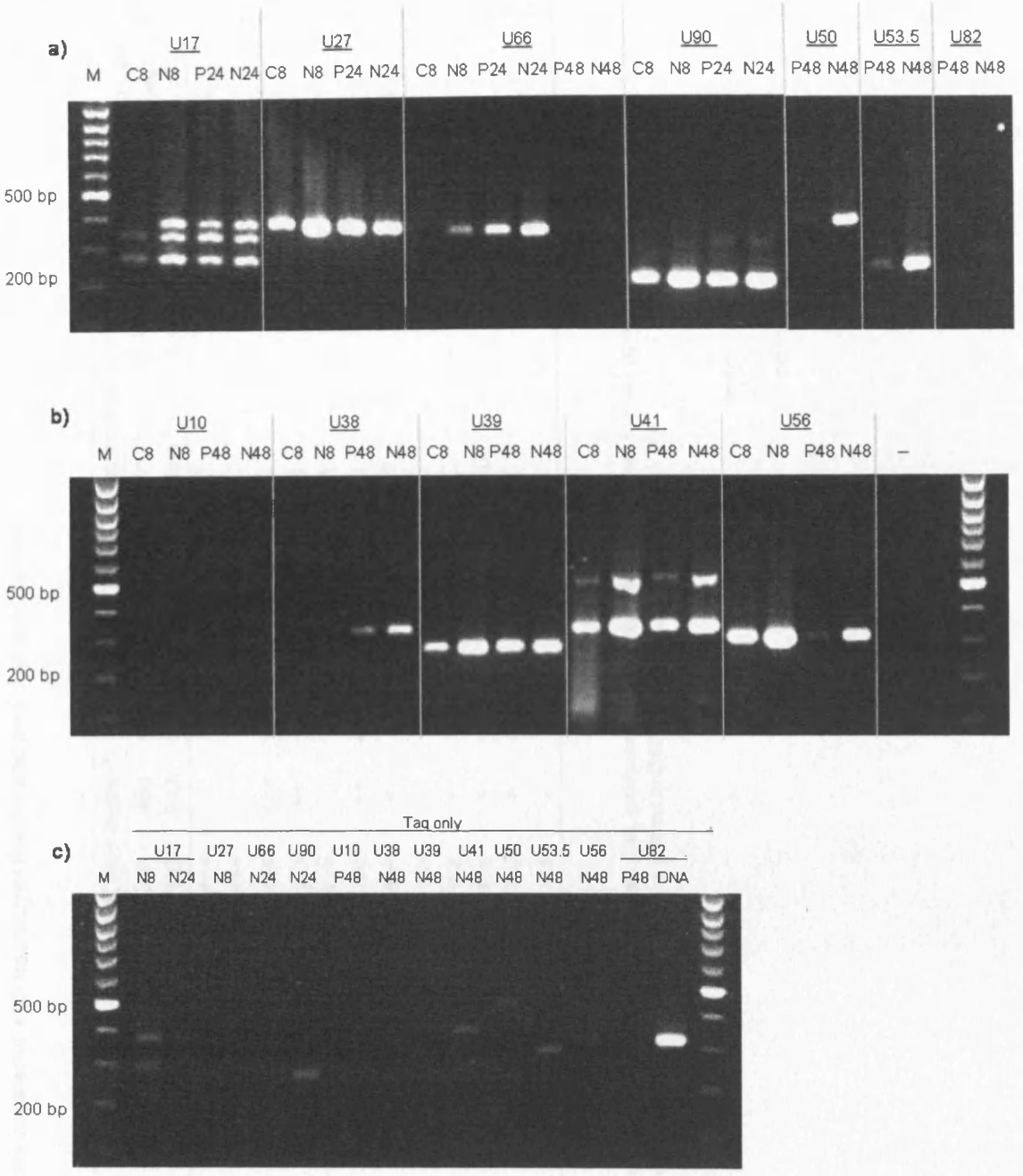


Figure 5.14 Agarose gel electrophoresis of RT-PCR products a-b) RT-PCR of aRNA from HHV-7 infected SupT1 cells using specific HHV-7 primers; c) Taq only, control PCR reaction without reverse transcription. RNA template: C8, CHX-treated 8hpi; N8, untreated 8hpi; P24, PAA-treated 24hpi; N24, untreated 24hpi; P48, PAA-treated 48hpi; N48, untreated 48hpi; DNA, HHV-7 DNA control; the 500bp fragment of DNA ladder contains 25ng DNA per lane.

Table 5.4 Temporal classification by RT-PCR results compared with microarray results and previous publication

HHV-7 ORF	Predicted function	RT-PCR (this thesis) [a]							Microarrays (this thesis)		RT-PCR with inhibitors [e]
		8h CHX	8h ND	24h PAA	24h ND	48h PAA	48h ND	Kinetic Class [b]	Transcription profiling [c]	PAA treatment & microarray [d]	
U17	IE-B, transactivator, US22 family	+	+++	+++	+++			E	B	IE/E	IE/E
U27	DNA polymerase processivity factor	++	+++	+++	+++			IE	B	IE/E	
U39	Envelope glycoprotein B (gB)	++	+++	n/a	n/a	+++	+++	IE	C	IE/E	IE
U41	Major DNA-binding protein,	++	+++	n/a	n/a	++	+++	IE	B	IE/E	IE
U90	IE-A transactivator	++	+++	+++	+++			IE	A	IE/E	IE
U38	DNA polymerase	-	-	n/a	n/a	++	+++	E	B	IE/E	
U66	DNA packaging,	-	++	+++	+++	+	+	E	B	IE/E	IE/L
U10	nuclear phosphoprotein	-	+	n/a	n/a	-	+	L	C	L	IE
U56	Capsid protein,	++	+++	n/a	n/a	+	++	L	C	L	
U50	Virion protein	n/a	n/a	n/a	n/a	+	+++	L	B/C	L	
U53.5	Major capsid scaffold protein	n/a	n/a	n/a	n/a	+	+++	L	C	L	
U82	Envelope glycoprotein L (gL)	n/a	n/a	n/a	n/a	-	+	L	C	L	

[a] Arbitrary scale of band intensity on agarose gel: +++, strong; ++, medium; +, weak; - negative; n/a, not tested; CHX, cycloheximide (200 µg/ml); PAA, phosphonoacetic acid (500 µg/ml)

[b] Kinetic class assigned based on RT-PCR results: IE, immediate-early gene unaffected by CHX; E, early gene inhibited by CHX but not PAA; L, late gene inhibited by PAA

[c] Transcription profiling of HHV-7 gene expression using microarrays without metabolic inhibitors

[d] Microarray analysis of HHV-7 gene expression in the presence of PAA

[e] Menegazzi et al., 1999

5.2.6 Host gene expression

The expression of cellular genes were analysed during HHV-7 infection of SupT1 *in vitro* using the custom-made microarrays. Clustering analysis showed that most host genes remained relatively unchanged throughout the time course (Figure 5.15). Any changes in cellular gene expression were of small magnitude relative to the mock infected cells (less than two-fold). Two clusters of genes can be tentatively identified (Figure 5.15). However, because of the small dynamic changes in gene expression, the significance of these genes was difficult to determine. Therefore to robustly identify differentially regulated genes, Significance Analysis of Microarrays (SAM) was used to compare cellular gene expression between mock-infected cells and HHV-7-infected cells at early (8-16 hpi) and late time points (48 and 72 hpi), and to identify significantly differentiated expression, using at a false discovery rate of 3%. Compared to mock-infected cells, 22 and 17 cellular genes were found to be significantly upregulated at early (8-16 hpi) and late (48-72) times post-infection, respectively, by HHV-7, whereas 12 and 29 cellular genes were found to be significantly downregulated at early and late times, respectively, (Figure 5.16). However, only one of the upregulated genes and 7 of the downregulated genes at late times had a fold-change above +1.5 or below -1.5. At early times only one gene had elevated expression levels above a fold change of +1.5 (Table 5.5). Some of these genes are involved in biological processes that have been identified in previous microarray studies of viral infections. For example, cholesterol biosynthesis was induced by HIV (van 't Wout et al., 2003). The expression levels of many genes involved in interferon-mediated response, apoptosis, protein synthesis, cytokine production and signalling, are increased by various viruses (Table 1.4). Changes in genes involved in transcription and cell cycle regulation have also been described previously (Table 1.4). Interestingly, hydroxysteroid (17-beta) dehydrogenase 12 was slightly upregulated in HHV-7 infected cells. Although the role of this gene in HHV-7 infection is undetermined, a related enzyme (3- beta-hydroxysteroid dehydrogenase) is encoded by vaccinia virus and may play a role in inhibiting host inflammatory response (Reading et al., 2003). However, in general we observed rather small fold-changes in the expression of only a small number of genes. in most published array studies of cellular trancription during viral infection, usually a larger number (up to hundreds) of host genes were up- and downregulated and very often

with much larger fold-changes. In comparison, our microarray data showed HHV-7 induced very little change in host gene expression.

Table 5.5 Significantly diffrentiated cellular genes in HHV-7 infected SupT1 cells

Gene	GenBank Accession	Mean fold change	Function
Upregulated at early times			
Programmed cell death 1 ligand 1	AI202996	+1.58	Apoptosis
Ubiquitin-conjugating enzyme E2A (RAD6 homolog)	AA045751	+1.02	DNA repair
Ankyrin repeat and SOCS box-containing 3	H40739	+1.00	Protein synthesis
Killer cell lectin-like receptor subfamily D, member 1	R78286	+1.00	NK cells and T- cells cytotoxicity
Hydroxysteroid (17-beta) dehydrogenase 12	N95053	+0.93	unknown
Interleukin 11	AI079412	+0.92	Hematopoiesis
Apoptosis antagonizing transcription factor	AI439571	+0.91	Apoptosis
Diacylglycerol kinase_ gamma 90kDa	BE468027	+0.89	Lipid metabolism
Transcription elongation factor A (SII), 1	AI421762	+0.85	Transcription
S100 calcium binding protein A3	N68995	+0.84	Cell differentiation
KIAA0053 gene product	AA262297	+0.84	GTPase activator
Joined to JAZF1	AI378171	+0.83	Transcription
v-ros UR2 sarcoma virus oncogene homolog 1	W24201	+0.83	Tyrosine kinase signalling
Regulator of G-protein signaling 14	AA428897	+0.82	G-protein signaling
Aldehyde dehydrogenase 9 family, member A1	AA010649	+0.82	Metabolism
Homeo box A10	AI187278	+0.78	Transcription
BCL2 binding component 3	AI821703	+0.78	Apoptosis
Leukotriene A4 hydrolase	AA987920	+0.77	Biosynthesis
Serine (or cysteine) proteinase inhibitor, clade D, member 1	H72143	+0.77	Peptidase activity
Centaurin_ beta 2	AA237088	+0.76	GTPase activator
Chromodomain helicase DNA binding protein 2	N90649	+0.73	Transcription
Associated molecule with the SH3 domain of STAM (AMSH) like protein	W95888	+0.70	Protein synthesis
Upregulated at late times			
Pirin	N29255	+1.98	Transcription
Associated molecule with the SH3 domain of STAM (AMSH) like protein	W95888	+1.46	Protein synthesis
Transcription elongation factor A (SII)	AI421762	+1.28	Transcription
SWI/SNF related_ matrix associated, actin dependent regulator of chromatin, subfamily a, member 3	BF51373	+1.28	Chromatin maintenance
Ankyrin repeat and SOCS box-coontaining 3	H40739	+1.19	Protein synthesis
Hydroxysteroid (17-beta) dehydrogenase 12	N95053	+1.13	unknown
Programmed cell death 1 ligand	AI202996	+1.12	Apoptosis
Succinate dehydrogenase complex	AA281913	+1.00	Carbohydrate metabolism
CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	AA418709	+0.99	Signal transduction, cell adhesion
Centaurin_ beta 2	AA237088	+0.97	GTPase activator
Mitogen-activated protein kinase	AA236814	+0.91	Heat shock
Signal recognition particle receptor ('docking protein')	AA283840	+0.90	Protein trafficking
Jun D proto-oncogene	AI339181	+0.88	Cell cycle
Isocitrate dehydrogenase 3 (NAD+) beta	AI566177	+0.85	Carbohydrate metabolism
TAF1 RNA polymerase II_ TATA box binding protein (TBP)-associated factor	AI806456	+0.83	Transcription
Macrophage stimulating_ pseudogene 9	AA262037	+0.81	
19A24 protein	BE46767	+0.72	
Downregulated at early times			
60s Acidic Ribosomal Protein P0		-1.43	Translation
Cytochrome P450_ family 11A1	AA142879	-1.31	Cholesterol metabolism
AA400583	AA400583	-1.29	
40s ribosomal protein s18		-1.16	Ribosomal RNA, protein biosynthesis
40s ribosomal protein s3A		-1.15	Ribosomal RNA, protein biosynthesis
Glucocorticoid receptor alpha mRNA	AA736674	-1.03	Transcription
Human triosephosphate isomerase		-0.99	Metabolism
Integrin_ alpha 2 (CD49B_ alpha 2 subunit of VLA-2 receptor)	AI538195	-0.92	Cell adhesion

GS2 gene	AI290788	-0.86	
Low density lipoprotein recept (familial hypercholesterolemia)	AA621897	-0.82	Cholesterol metabolism
Leukotriene A4 hydrolase	AI521212	-0.81	Biosynthesis
Homeo box A5	AI223317	-0.73	Transcription
Downregulated at late times			
Cytochrome P450 11A1	AA142879	-3.10	Cholesterol metabolism
60s Acidic Ribosomal Protein P0		-2.39	Translation
40s ribosomal protein s18		-2.10	Ribosomal RNA
Histone 1_ h2bk	AA280257	-2.09	Histone
Guanine nucleotide-binding protein		-1.58	G-protein signaling
Branched chain keto acid dehydrogenase kinase	W78018	-1.58	Protein synthesis
Human triosephosphate isomerase		-1.55	Metabolism
Death associated protein 3	AA858184	-1.25	Interferon-mediated inflammation
Peptidyl-prolyl:cis-trans isomerse		-1.23	
Heterogeneous nuclear ribonucleoprotein C	AI096954	-1.21	Ribonucleosomes
Lymphoid enhancer-binding factor 1	AF086339	-1.19	
TATA box binding protein	AA470112	-1.19	Transcription factor
Tubulin gaunachain		-1.12	
Aa459022	AA459022	-1.09	n/a
Low density lipoprotein recept (familial hypercholesterolemia)	AA621897	-1.07	Cholesterol metabolism
Peroxiredoxin 6	AA016074	-1.05	Oxidative Stress
Human ribosomal protein L30		-1.01	Protein synthesis, mitochondrial
Interleukin 18 (interferon-gamma-inducing factor)	AI394321	-1.01	Interferon-mediated inflammation
Muscleblind-like 3 (Drosophila	AA460747	-0.99	
UL16 binding protein 1	AI830832	-0.98	Cytokine signaling
Homo sapiens transcribed sequeuence	AA748610	-0.95	Cell cycle
Ai245910	AI245910	-0.94	
Tripartite motif-containing 29	AA293775	-0.93	
Aw449304	AW449304	-0.92	
SWI/SNF related_ matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	AI279006	-0.92	Transcription regulation
Postmeiotic segregation increased 2-like 5	AI015925	-0.82	
AT-hook transcription factor A	AA904466	-0.82	
SH2 domain protein 2A	AA613070	-0.80	Signal transduction
C-type lectin, superfamily member 12	AA478515	-0.74	Cell adhesion

* Mean $\log_2$ fold change of the triplicate spots on custom-made host-pathogen arrays

There were three possible reasons for this lack of change in host gene expression during HHV-7 infection. Firstly, the relatively small number of cellular probes (approximately 1000) on the custom-made array may have meant that genes that underwent transcriptional changes were not represented on the array. However, this was not likely, because when the same samples were hybridised onto the Human cDNA Array v.2 (Human Genome Mapping Project, MRC, UK), which contained approximately 11,000 human gene probes, the clustering results were similar to the that obtained using custom-made arrays. Secondly, sense strand RNA amplification, which favoured the 5'-end of mRNA transcripts, may have introduced a bias that prevented the detection of subtle changes in cellular transcription. However, the signals of the host gene probes were sufficient for detection, and the pass-rates were adequate (Figure 5.17). Therefore sense-strand RNA amplification had produced enough cellular RNA for hybridisation. Sense-strand amplified RNAs have been analysed by quantitative RT-PCR and was shown to preserve known expression ratios (Goff et al., 2004) and the relative changes in viral transcription levels were clearly conserved after amplification (section 5.2.3). However, it was unknown whether the differences in cellular mRNA abundance were affected by sense strand mRNA amplification in this study. If this was true, however, it would be very difficult to analyse host gene expression during HHV-7 infection using microarrays, because RNA amplification was necessary to ensure a high MOI and synchronous infection.

A third possible reason for the lack of differential expression may be that HHV-7 caused no major changes in the expression of host genes during infection of SupT1 cells *in vitro*. This was also difficult to comprehend: how the SupT1 cells could show no response to a virus that infected 80% of the culture population (section 5.2.1), replicating its genome and producing virions (section 5.2.3). However, if true this may reveal details about the ability of the HHV-7 virus to infect 90% of the world's population without apparent pathologies (section 1.1.9) and would represent the extreme of herpesvirus mutualism with human hosts. So far this phenomenon has not been observed in any other virus. All microarray studies of viral infections have shown that host gene expression was extensively altered during infection (section 1.2.4.3, Table 1.4). It will be interesting to study, if true, why HHV-7 induced no large changes in cellular transcription.

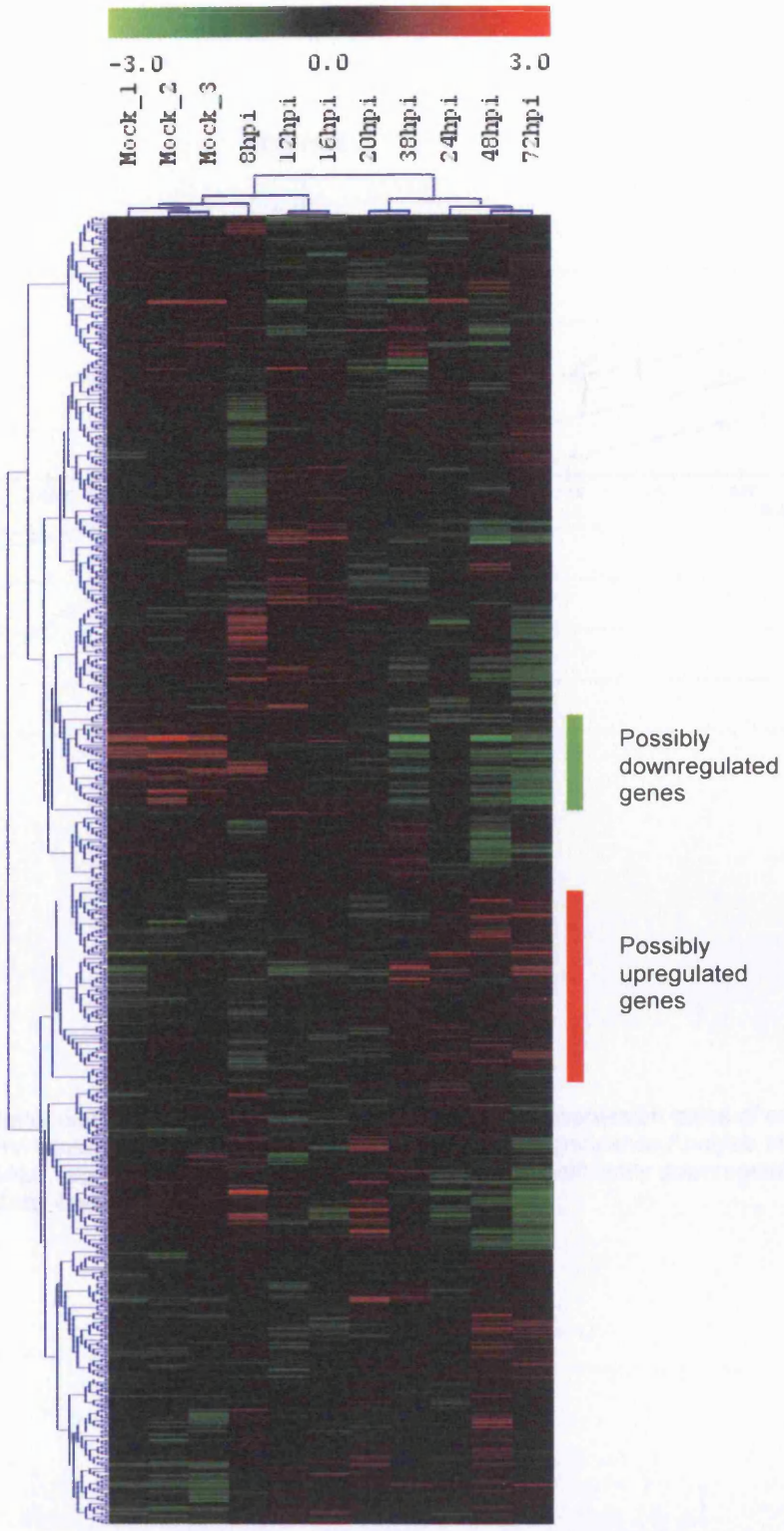


Figure 5.15 Cluster analysis of host gene expression in SupT1 cells mock-infected and infected with HHV-7 over a 72-hour period, measured using custom-made microarrays, with two clusters representing genes possibly upregulated (red rectangle) and downregulated (green rectangle) by HHV-7

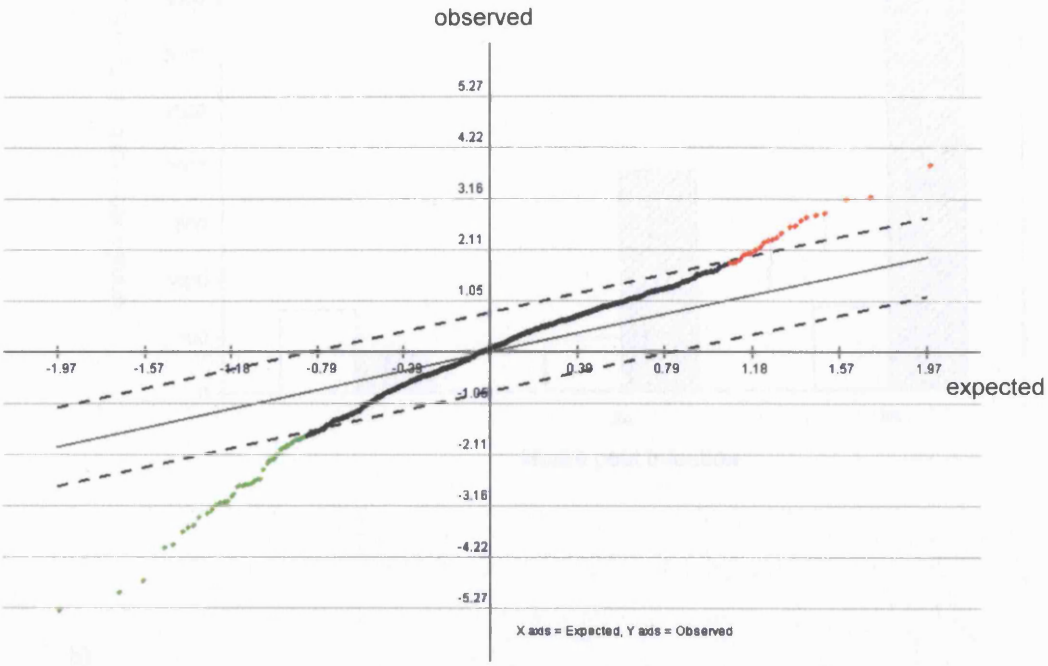
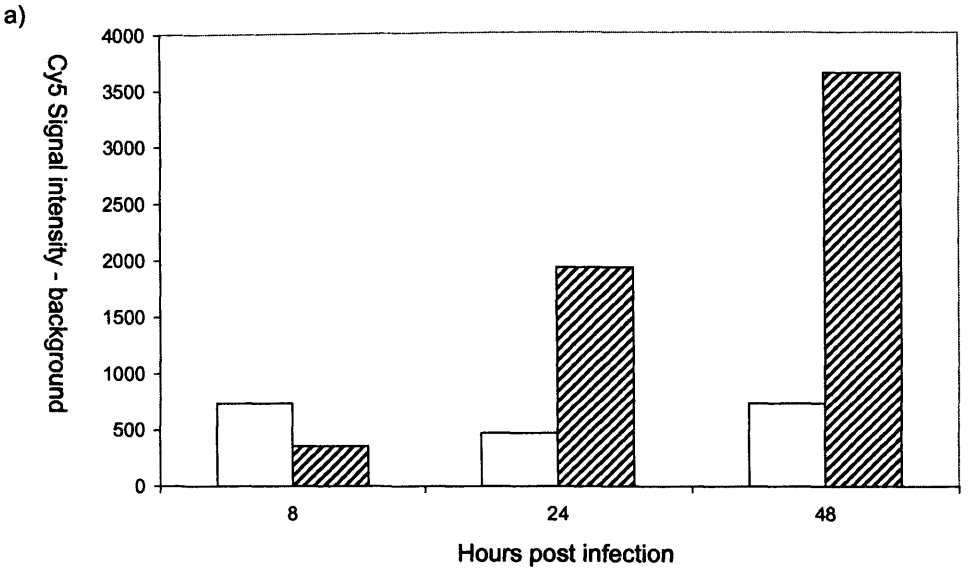


Figure 5.16 Scatter plot of observed (y-axis) vs expected (x-axis) expression ratios of cellular genes from HHV-7 infected cells and mock-infected cells using Significance Analysis of microarrays (SAM). Significantly upregulated genes (red dots), significantly downregulated genes (green dots) and unchanged genes (black dots)



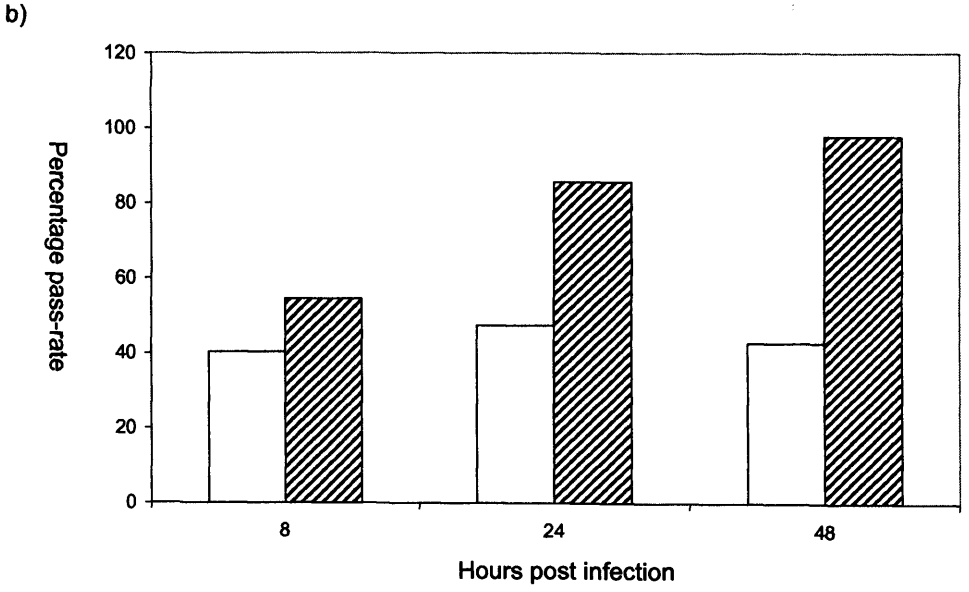


Figure 5.17 a) Background-subtracted Cy5 intensities and b) percentage pass rates of cellular probes (white bars) and HHV-7 probes (striped bars) on the custom-made microarrays hybridised to amplified RNA extracted from HHV-7-infected SupT1 cells at various time points during infection.

5.3 Summary and Discussion

This chapter described the analysis of HHV-7 gene expression using custom-made microarrays, demonstrating the power of microarrays as a large-scale analysis tool capable of deciphering the whole transcriptome of the virus. Transcription profiling revealed the coordinated expression of viral genes during lytic infection of T-cells *in vitro*. The regulated induction of HHV-7 genes from 8 to 72 hours post infection reflected the sequential stages of viral infection cycle, consistent with the classic model of temporal regulation of herpesvirus gene expression. From the profiling data we have also estimated the HHV-7 lytic cycle to be between 36 to 60 hours long, predicted functions for previously uncharacterised genes, and identified the possible roles of certain transcription regulators. Microarrays also enabled the assignment of 68 HHV-7 genes into temporal classes A, B, and C (corresponding to IE, E and L). A final consensus of temporal classes (Figure 5.18) was achieved from the results of transcription profiling (section 5.2.3); inhibitor treatment coupled with microarrays analysis (section 5.2.4) and RT-PCR (section 5.2.5). Genes of the same temporal class were sometimes adjacently situated in the viral genome, suggesting the possibility of polycistronic transcription (Figure 5.18).

Temporal classification of the HHV-7 genes varied between the different methods. This was probably due to the different features of the dataset emphasised in the different methods. The transcription profiling used (section 5.2.3) was a comparison approach, measuring the relative changes in expression over time. It revealed the temporal regulated patterns of transcription in the absence of metabolic inhibitors, therefore revealing viral gene expression with minimal artificial intervention. The use of microarrays coupled with inhibitor treatment (section 5.2.4) was a rapid method for identifying statistically significant genes that were differentially expressed between treated and untreated cells. The microarray technique was not very sensitive for the detection of low abundance transcripts, especially in 8 hpi samples, therefore was unable to distinguish between IE and E genes. Nevertheless, its high throughput nature allowed many genes to be analysed simultaneously and rapidly. Low abundance genes that require further investigation using other methods was rapidly identified for follow up in future experiments.

RT-PCR, on the other hand, was a highly sensitive qualitative approach capable of detecting very low abundance mRNAs (section 5.2.5). It provided a simple positive/negative answer to whether a gene was transcribed in the presence of metabolic inhibitors. However, it has been suggested that 35 PCR cycles were too sensitive and may lead to detection of leaky E transcription under IE conditions and misassignment of kinetic classes (Pellett, 2001). For example, U42 and U73 of HHV-6, classified as IE genes by qualitative RT-PCR (Mirandola et al., 1998), were found to be dependent on *de novo* protein synthesis by quantitative RT-PCR (Rapp et al., 2000). This may explain the inconsistencies between microarray-based methods and RT-PCR in the classification of IE and E genes (Table 5.3 – 5.4).

In summary, RT-PCR was more sensitive for low abundance transcripts and has the advantage of revealing transcription kinetics for different splice forms. Microarrays offered faster, larger scale temporal classification of herpesvirus genes in the presence or absence of toxic metabolic inhibitors. However, the strength of microarray technology was its capacity to disclose the relative changes in gene expression, resulting in a global picture of the molecular events during HHV-7 infection of T-cells.

The microarray data also showed that HHV-7 infection induced very little changes in host gene expression, with only 8 genes showing significantly differentiated expression above a fold change of 1.5 (section 5.2.6). However, further experiments are required to validate this finding.

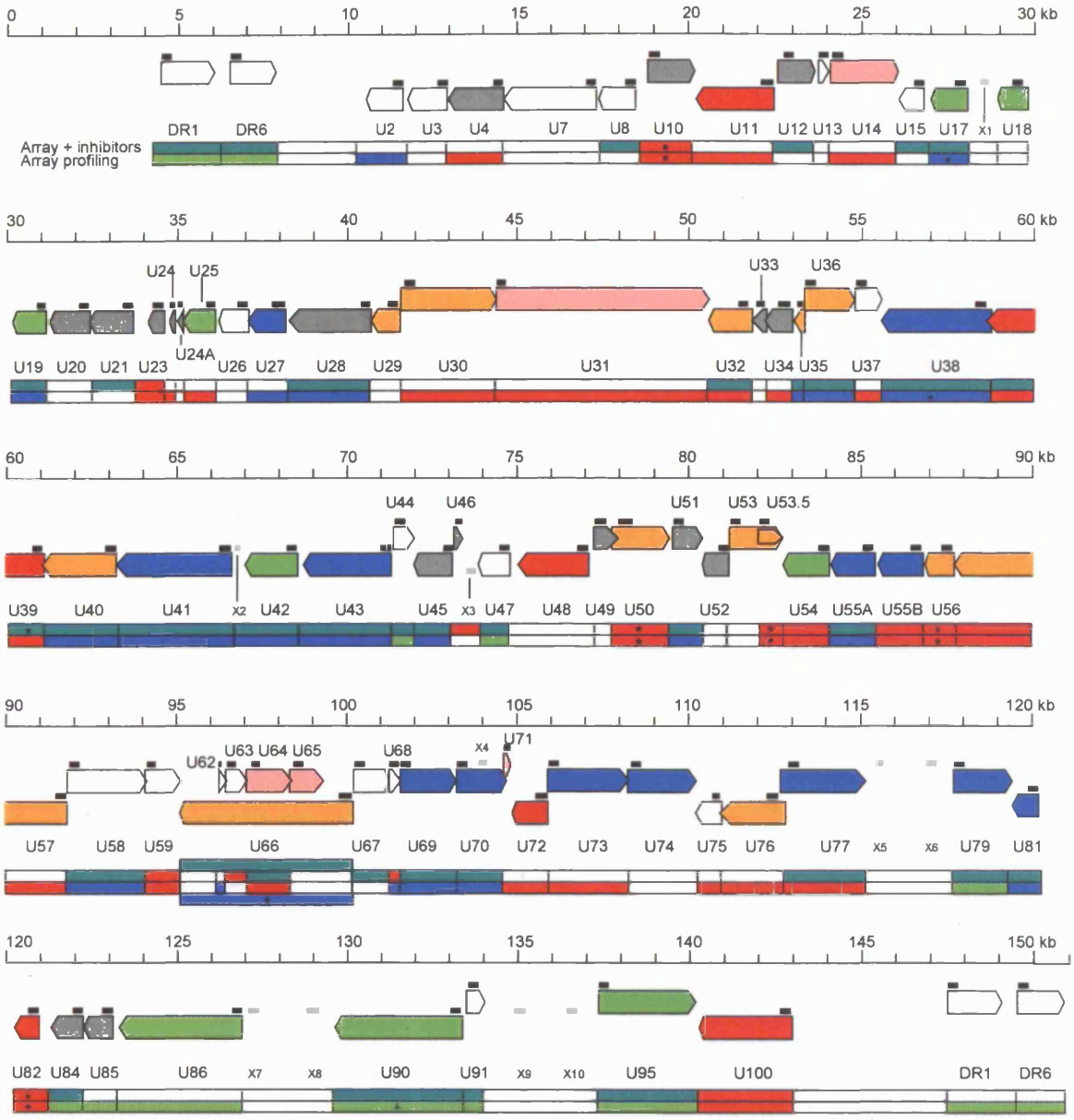


Figure 5.18 Temporal map of HHV-7 genome (RK strain). Predicted functions of the ORFs are colour-coded:

transactivators; replication functions; capsid & assembly; tegument; envelop; othe; unknown.

Array probes of approximately 300 bp were designed for each ORF (small black rectangles); also for ten intergenic regions (small grey rectangles) to measure the expression level of their corresponding ORF. HHV-7 ORFs were classified as IE (green block), E (blue block) or L (red block) genes by expression profiling using arrays; also as IE/E (turquoise block) or L (red block) by arrays with inhibitor (PAA). Kinetic classification of some ORFs were confirmed by RT-PCR (marked with *).

CHAPTER 6

CONCLUSIONS & FUTURE DIRECTIONS

6.1 HHV-7 gene expression

This thesis described the production of a sensitive, specific and reproducible host-pathogen microarray for the detection of human herpesvirus type-7 (HHV-7) mRNAs, and its application in expression profiling of the viral transcriptome during human T-cell infection *in vitro*. Analysis of the microarray results has revealed that HHV-7 genes were co-ordinately expressed in an ordered pattern, which reflected the sequential stages of viral lytic replication cycle (transcription activation, replication, virion assembly) that is common amongst herpesviruses (Pellett, 2001; Roizman, 2001a; Roizman, 2001b). HHV-7 genes were classified into three temporal classes A, B or C in the absence of metabolic inhibitors, therefore represented viral lytic infection without any external intervention. These classes generally mirrored the herpesvirus immediate-early (IE), early (E) and late (L) genes, respectively. This correlated well with previous assignments of herpesvirus genes into their temporal classes by transcription profiling (Jenner et al., 2001; Paulose-Murphy et al., 2001). A selection of class B and C genes have been confirmed to be E and L genes, respectively, by the use of phosphonoacetic acid (PAA) to block viral DNA replication, hence L gene expression. The confirmation of HHV-7 IE classification, however, would require further work, due to cellular toxicity of cycloheximide (CHX) at 16 hours post infection (hpi) and low expression level of early genes at 8 hpi. A useful experiment to do would be to compare HHV-7 gene expression at 12 hpi in the presence and absence of cycloheximide, in order to distinguish between viral IE and E genes.

Nonetheless, we have shown that transcription profiling using microarrays allowed the temporal classification of herpesvirus genes consistent with metabolic inhibitor assays. Microarray analyses offer higher throughput, reproducibility, speed and a means to measure genome-wide viral gene expression in their 'natural' state, thus have potential advantages over metabolic inhibitor treatment assays. The HHV-7 host-pathogen microarrays we created could be used to study the transcriptome of clinical HHV-7 isolates from latently-infected adults, or from children with primary infection, and these can be compared with the results obtained in this study using the DC strain.

Being the third member of the human betaherpesviridae family, unlike human cytomegalovirus, HCMV (Chambers et al., 1999) and human herpesvirus type-6, HHV-6 (Ohyashiki et al., 2005), no extensive viral gene expression studies have been performed on HHV-7. This thesis has made a step to complete that picture, showing that HHV-7 gene expression patterns were similar to the other betaherpesviruses, and that the three human betaherpesviruses all follow similar lytic replication events *in vitro*. However, the genomes of HCMV and HHV-6 contain extra ORFs that are not found in HHV-7, and it would be interesting to compare the gene expression profiles between these betaherpesviruses. This could be performed using the custom-made microarrays, as array probes for HHV-6 ORFs and a subset of HCMV genes have been generated in our laboratory as separate projects.

In addition, findings in this study also supported the predicted functions of HHV-7 ORFs (Megaw et al., 1998; Nicholas, 1996), and suggested functional classes of some previously uncharacterised genes. We have also estimated the time of action of some transcriptional regulators, hence indicating their possible roles in particular stages of the viral lytic infection cycle. HHV-7 U42 and U54 may be responsible for regulating the transcription of genes encoding enzymes for viral DNA replication and structural proteins, respectively. The functional significance of these two transactivators could be examined by gene knock out experiments, using gene silencing techniques such as RNAi (Hannon, 2002; Kuwabara and Coulson, 2000). RNAi has been used to study host-pathogen interactions in bacterial (Agaisse et al., 2005; Walduck et al., 2004) and viral infections (Lieberman et al., 2003; Wang et al., 2003). The impacts on viral DNA replication and virion assembly upon RNAi silencing of HHV-7 U42 and U54 may provide useful information on the functions of these transcriptional regulators, and thus improve understanding of how herpesviruses regulate their replication events during infection.

We have demonstrated the power of microarray transcription profiling in the study of large genome viruses. So far these techniques have only concentrated on herpesviruses, while no global gene expression profiling studies have been published for the other virus families with large dsDNA genomes, such as the poxviridae and the adenoviridae. Gene expression in poxviruses are also temporally regulated (Broyles, 2003), thus is particularly suited for transcriptional profiling using

microarrays. Although the genome of vaccinia virus has been fully sequenced (Goebel et al., 1990) and a microarray has been generated for vaccinia genes (Wenli et al., 2004), it has not yet been used for the transcriptome profiling of the vaccinia virus.

6.2 HHV-7 culturing

This thesis has also documented the optimisation of HHV-7 quantification methods and protocols for maximising the infectious titre of the virus in SupT1 cells *in vitro*. Previous attempts to improve HHV-7 infectivity by methods such as ultracentrifugation (Ablashi et al., 1998) were without success. We have shown that growing HHV-7 in small volumes was an efficient method for maximising the infectious titre of the virus stocks harvested (section 4.2.4.1). This is also the first report on the effects of spinoculation and polybrene on HHV-7 inoculation efficiency. The multiplicity of infection (MOI) of HHV-7 in T-cells *in vitro* was reduced in the presence of polybrene (8 μg/ml), but was improved by spinoculation at 1000 g, although the increase was not as high as for some other viruses (Pietroboni et al., 1989). Therefore the effects of spinoculation and polybrene on HHV-7 infectivity were similar to that on its closest relative HHV-6 (Black et al., 1989), but not to the more distant gammaherpesvirus HHV-8 (Inoue et al., 2003), of which the infectivity was increased by both spinoculation and polybrene. These findings would benefit future studies requiring the infection by cell-free HHV-7 virus at a high MOI, particularly gene expression analysis, which has been hindered by the low infectivity of HHV-7 (Pellett, 2001).

RNA amplification has been applied to analyse limited clinical samples such as those extracted by laser capture microdissection or flow cytometry-sorted cells. We have demonstrated that the RNA amplification technique could also be used to amplify RNA from low cell numbers in order to maximise the MOI in virology studies. This may be useful for other viruses with low infectivity-to-particle ratios in cell-free virus form, for example human T-cell leukaemia virus type 1 (HTLV-1) (Sasada et al., 2005), human and murine cytomegaloviruses (Ho et al., 1993; Hodgkin et al., 1988; Hudson, 1988; Osborn and Walker, 1968), and retroviruses (Bahnson et al., 1995;

Chuck et al., 1996; Ho et al., 1993; Kotani et al., 1994). This technique may also benefit gene expression studies on slow growing micro organisms such as *Plasmodium falciparum*, and *Mycobacterium tuberculosis*. However RNA amplification adds cost and time to the experiment therefore may not be suitable for some applications.

6.3 Host genes expression by HHV-7

We have found little change in the expression of cellular genes during HHV-7 infection of SupT1 cells *in vitro*. This could be due to technical limitations or HHV-7 may be very efficient at evading host response. Confirmation requires further work on validating the use of sense-strand RNA amplification technique on cellular mRNAs. The sense-strand RNA amplification method used in this thesis can be used to amplify RNA from samples that are known to contain significantly different amounts of certain species of cellular mRNA. However, if HHV-7 truly induces so little response in the host during infection, more analysis on how HHV-7 achieves this may provide very useful information about host-pathogen interactions and viral evasion of host immune response. In particular, silencing of candidate HHV-7 genes may give clues to which, if any, of the HHV-7 genes are responsible for the virus's ability to establish such silent infection in the host cell. Moreover, this naturally non-pathogenic virus may be a very useful potential gene therapy vector. However, before HHV-7 can be developed into a gene therapy vector, more studies are required to characterise HHV-7 cell tropism, also the mechanisms for cell attachment, entry and expression of viral-encoded antigens on the cell surface.

Appendix A – Sequence alignments of HHV-7 array probes

HHV7DR1 & DR1'

```
HHV7_RK_DR1       ATGACTGCTGCAACCACAGAACATTTTGCTCTCCGCGCGGCACTCAATCGTTACTGGTGG 60
HHV7_RK_DR1'      ATGACTGCTGCAACCACAGAACATTTTGCTCTCCGCGCGGCACTCAATCGTTACTGGTGG 60
H1_020206         ------------------GAACATTTTGCTCTCGTCGCGGCACTCAATCGTTACTGGTGG 42
A1_020228         ------------------------------------------------------------

HHV7_RK_DR1       CTGCTTCTGGGACGACACAAGCTCAGTTTGGTATGCAACT-ACGTCACAG-CTCATCGCC 118
HHV7_RK_DR1'      CTGCTTCTGGGACGACACAAGCTCAGTTTGGTATGCAACT-ACGTCACAG-CTCATCGCC 118
H1_020206         CTGCTTCTGGGACGACACAAGCTCANNTTGGTATGCAACT-ACGTCACAG-CTCATCGCC 100
A1_020228         ------------------AAGC-CAGTTTGGTATGCNACTGACGTCACAGGCTCCTCGCC 41
                                    **** **  ********* *** ********* *** *****

HHV7_RK_DR1       AAC-AGTTACTGCCGCTGCCGTGGCCCGAACAGGAATTTCTCCAACTTGA-CCCGGCCCC 176
HHV7_RK_DR1'      AAC-AGTTACTGCCGCTGCCGTGGCCCGAACAGGAATTTCTCCAACTTGA-CCCGGCCCC 176
H1_020206         AAC-AGTTACTGCCGCTGCCGTGGCCCGAACAGGAATTTCTCCAACTTGA-CCCGGCCCC 158
A1_020228         AACCAGTTACTGCCGCTGCCGTGGCCCGAACAGGAATTTCTCCAACTTGAACCCGGCCCC 101
                  *** ********************************************* *********

HHV7_RK_DR1       CTACTCCAATCTCCGCAACCGT-GTCGC-TCACCATCTCCATCGCGGCTGGCCAGCGGCA 234
HHV7_RK_DR1'      CTACTCCAATCTCCGCAACCGT-GTCGC-TCACCATCTCCATCGCGGCTGGCCAGCGGCA 234
H1_020206         CTACTCCAATCTCCGCAACCGTTGTCGCATCACCATCTCCATCGCGGCTGGCCAGCGGCA 218
A1_020228         CTACTCCAATTTCCGNCNCCCCCGT-GTACTCCCATCTCCNTTNCGGGTGG--------- 151
                  ********** ****   **   ** *     ******** *  *** ***

HHV7_RK_DR1       CA-CAACACAtgtaagctaccgtacatctctttca-caaacccaaggctcacatagagac 292
HHV7_RK_DR1'      CA-CAACACATGTAAGCTACCGTACATCTCTTTCA-CAAACCCAAGGCTCACATAGAGAC 292
H1_020206         CNACAACTCATGTAAGCTACCGTACATCTCTTTCATCAAACCCAGGGCTCACATAAAGAC 278
A1_020228         ------------------------------------------------------------

HHV7_RK_DR1       aagcacaagctcgcgcaatgacattaaaacctcccatcattgtcctttcctgtcgctttg 352
HHV7_RK_DR1'      AAGCACAAGCTCGCGCAATGACATTAAAACCTCCCATCATTGTCCTTTCCTGTCGCTTTG 352
H1_020206         AAGCACAAGCTCG--CAATCACT---------------------------------------- 299
A1_020228         ------------------------------------------------------------
```

HHV7DR6 & DR6'

```
HHV7_RK_DR6       ATGAGTGCAGAAATGCTCCGCGCTGTTCAGCTCCAGCCAAGACGCCGGGGACATTCCTCA 60
HHV7_RK_DR6'      ATGAGTGCAGAAATGCTCCGCGCTGTTCAGCTCCAGCCAAGACGCCGGGGACATTCCTCA 60
H2_020425         ---------------------TCTGTTCAGCTCCAGCCAAGACGCCGGGGACATTCCTCA 39
A2_020206         AGTGATTAGANAATGCTCCGCGCTGTTCAGCTCCAGCCAAGACGCCGGGGACATTCCTCA 60
                                      ************************************

HHV7_RK_DR6       TCTCCCACTTCCCCTCCACTCG------AAGGAGAGCCCAGTCCCAAGAGACTCCAATCG 114
HHV7_RK_DR6'      TCTCCCACTTCCCCTCCACTCG------AAGGAGAGCCCAGTCCCAAGAGACTCCAATCG 114
H2_020425         TCTCCCACTTCCCCTCCACTCG------AAGGAGAGCCCAGTCCCAAGAGACTCCAATCG 93
A2_020206         TCTCCCACTTCCCTNTCGNANNTTCGTNNAGGAGAGCCCAGTCCCAAGAGACTCCAATCG 120
                  *************   *            *********************************

HHV7_RK_DR6       AGCAACAGTCACCAAGGGCGTAGAGGCAGACCTAAACCCAGAGCTAAAACATGGAGCGAA 174
HHV7_RK_DR6'      AGCAACAGTCACCAAGGGCGTAGAGGCAGACCTAAACCCAGAGCTAAAACATGGAGCGAA 174
H2_020425         AGCGACAGTCACCAAGGGCGTAGAGGCAGACCTAAACCCAGAGCTAAAACATGGAGCGAA 153
A2_020206         AGCGACAGTCACCAAGGGCGTAGAGGCAGACCTAAACCCAGAGCTAAAACATGGAGCGAA 180
                  *** ************************************************************

HHV7_RK_DR6       GCTTTATCCCACCGGTCCTTCCTCAACATTTACGCGTGGCTGTCTTTGAGT-CGAGGGTC 233
HHV7_RK_DR6'      GCTTTATCCCACCGGTCCTTCCTCAACATTTACGCGTGGCTGTCTTTGAGT-CGAGGGTC 233
H2_020425         GCTTTATCCCACCGGTCCTTCCTCAACATTTACGCGTGGCTGTCTTTGAGT-CGAGGGTC 212
A2_020206         GCTTTATCCCACCGGTCCTTCCTCAACATTTACGCGTGGCTGTCTTTGAGTGCGAGGGTC 240
                  ************************************************** ********

HHV7_RK_DR6       TCCGCGAAAAGTGTACGGATATGCCTTCAGGCACAGA-GGAGAACTCGTAGCATTG-CCA 291
HHV7_RK_DR6'      TCCGCGAAAAGTGTACGGATATGCCTTCAGGCACAGA-GGAGAACTCGTAGCATTG-CCA 291
H2_020425         TCCGCGAAAAGTGTACGGATATGCCTTCAGGCACAGA-GGAGAACTCGTAGCATTG-CCA 270
A2_020206         TCCGCGAAAAGTGTACGGATATGCCTTCAGGCACAGNAGGAGAACTCGTAGCATTGANCA 300
                  ************************************  ******************  **

HHV7_RK_DR6       TGGCCGCCTAACTGGAGCCTGGAACTTCACCACGATCCCTATCGAGACGCCAGAGCACAA 351
HHV7_RK_DR6'      TGGCCGCCTAACTGGAGCCTGGAACTTCACCACGATCCCTATCGAGACGCCAGAGCACAA 351
H2_020425         TGGCCGCCTAACTGGAGCCTGGAACTTCA---------------------------------- 299
A2_020206         TGGCCGCCTAACTG-------------------------------------------------- 314
                  **************
```

HHV7U2

```
HHV7_RK_U2      TCTGCGCAGACCACGCTTCCAGAGACCATAAGGGGTGTCCGCAACTCTGTAAATGGCGTT 660
A3_020419       ------------------------------------------------------GGGGTT 6
                                                                     ** ***

HHV7_RK_U2      GTCCACGGAACGGTTAAATGCATAAAATTCACCCCGATCTCCAATAAGCATCAAAGCAGA 720
A3_020419       -TANANCGTTTGNCNAGAANCTTNCGGTCCTTCNCGGAAN---ANTGGCNNCAAGACATA 62
                 *  *  *   *  * *  * *   * *  * **       *   **  ***  ** *

HHV7_RK_U2      TAAATCAGAATAGAGCACCAGACCGATGACCCAAACTTCCACAAAAGGACAGCAGACAAA 780
A3_020419       CTCN-CGNAAGAGGCAAANCGCNTCAAGCTCGAGNNCGT---CGCGTGTAGAGGNCGCG 118
                     *  ** ** ***     *      * * *             *   **  * *

HHV7_RK_U2      ATTTCTGGTTAACTTGCAGACTTCGTCTTCCGAAATGCCAAGCATCTCCGAAGAGGCAAA 840
A3_020419       AACAGCGGAGGAGGGGGANAGGNTNNNGNN-GANNTGNGGTGTAGGNGNANGGANGTNTG 177
                *     **   *  * * *          **  **    * *        ** *

HHV7_RK_U2      CCTCAACTCAGCGCCGGAAGGCCACACGAGAGGAAGTCGATGGTGAGCATATTGCAAAAC 900
```

```
A3_020419       NCCNAGNGGAGANGCGANAGANNNN-TGNGNGGNTCGGGAAGNGGTGCGGTGCGNGGAGG 236
                 *  *    **   **  **       * * **     ** *  * * **     *   *

HHV7_RK_U2      AAAATGAGCAACGTCTCCGAAAGATACGGTTTCCGCAGCGAAGCTCTGC--AAAAACGTG 958
A3_020419       NGGAGGAGTAGCGGTTGTGTNCGCGGCAG-----GCGGAAGNNCTCTGTCAAAAANCGTG 291
                   * *** * **  *  *   *   * *     ** *     *****   **** ****

HHV7_RK_U2      GTCC-AGTAGACCTCCGGAACAGGTGATTCAACT--GTATTCGATTCCAGATT-AAACAA 1014
A3_020419       GNNCCAGTAGACCTCCGGAACAGGTGATTCAACNTGNTATTCGATTCCAGANTTAAACAA 351
                *  * ****************************    ************** * ******

HHV7_RK_U2      ACCGTTCGATGCATTCGACTCAACAGCCGTTTCCACACCGGAATCCAAAAG-CGGAGAAT 1073
A3_020419       ACCNTTCGATGCATTCGACTCAACAGCCGTTTCCACACCGGAATCCAAAAGTCGGAGAAT 411
                *** ************************************************ ********

HHV7_RK_U2      CTGACAT 1080
A3_020419       CTGAT-- 416
                ****


HHV7U3

HHV7_RK_U3      TCAGTCTCATACTCAAGTTCGTAAACAGTCGTCACAGTATGTTTTATGTAACCAAGAACA 780
A4_020413       ----------------------------------------NCNNGNTNNTNAANCGG 17
                                                         *      *

HHV7_RK_U3      GTAAGTTCTTCAGGACAACACACCAACGCGGCGTACTTTTTGAGCT---TATCTTCCTCA 837
A4_020413       NNAATNCCCCCGAGTCANCNCACGGGNATTNC-TGCTNCATGGNCCGCATACCGAACNGA 76
                  **   *  *  * ** * ***         * * **   **  *    ** *   *  *

HHV7_RK_U3      TATCCCTCCACACCA--TAAGTCAA--AACAAGTTCAAACCCCT-CTGGAAAAACCAAAG 892
A4_020413       GCGACGCNCATATCAACTCAGTNGATTAACAGTTANAAACCTNNACNGGAAAAACCAAAG 136
                    *   ** * **  * ***  *  ****  *  *****    * **************

HHV7_RK_U3      G-CACCCTCAAATGCGGATTCAGAAACAAAAACATTTTTAG--ACTATTAAAATTTGTAT 949
A4_020413       GACACCCTCAAATGCGGATTCAGAAACAAAAACATTTTTANGANCTATTAAAATTTGTAT 196
                * ****************************************    ****************

HHV7_RK_U3      AAACACGTAGAAAATCCACGAGTTTAGCCTTAGCAGATTCAAAGATTTTCTCTTTTCTTT 1009
A4_020413       AAACACGTAGAAAATCCACGAGTTTAGCCTTAGCAGATTCAAAGATTTTCTCTTTTCTTT 256
                ************************************************************

HHV7_RK_U3      TTTTATCA-GATAAAAACGTTTTCTCAACGTTGGCAACCTTCAAATCCTTTTTCCCGTCA 1068
A4_020413       TTTTATCACGATAAAAACGTTTTCTCAACGTTGGCAACCTTCAAATCCTTTTTCCCGTCA 316
                ******** ****************************************************

HHV7_RK_U3      ACATGAGC-AACATATTTTCGTATAAACGAATTTTCCTTTTTTAAATCTTTTGCCACAAC 1127
A4_020413       ACATGAGNCAACATATTTTCGTATAAACGAATTTTCCTTTTTTAAATCTTTTGCCACAAC 376
                *******  ****************************************************

HHV7_RK_U3      TCTCAAAACACGCTCTTCAAAATTGAAGTCAGACCCCAAATCTGCCAT 1175
A4_020413       TCTCAAAACACGCTCTTCAAAATTGAAGTCAGACCCCAAATCTGCA-- 422
                *********************************************


HHV7U4

HHV7_RK_U4      TGTAGCTTTAGTCAGAACTCGATAAAACGGAGAAGTGCACTCTTCAAACCTTAGACTTTG 1320
A5_020313       ----------------------------------------------GCAACNAATACCAGT 15
                                                               * *  * **

HHV7_RK_U4      AAGAAACGCGAGAAAACCGTCTAAAGTCATTGTTGCCTTGTCCCTCCTTTTAAATGAAAC 1380
A5_020313       AAAAAAAAAAACAAATNNAAAAAAAANNNGATCTGTGCCGGTNCC--CCTTAAAGTGANAC 73
                ** ***    * ***       ***     * *   *  ** **  * ** ** *** **

HHV7_RK_U4      GGAGTTTTTCTCCACTAAAACAGCCCATAGTAGACTGTCATCGTAAAATTTTAA---GCC 1437
A5_020313       GAGNNTCTCCNC----AAAACAGCCCATAGTAGACTGTGNTCGTAAAATNCNCACCAGCC 129
                *    * * * *    ********************  *********    *   ***

HHV7_RK_U4      TCGAGTTTGGAACAAATCTTCTGAAGTGAGATGACCGCTATCCGCAA-TCATCTCCTCTG 1496
A5_020313       TCGAGTTTGGAACAAATCTTCTGAAGTGAGATGACCGCTATCCGCANATCATCTCCTCTG 189
                ************************************************  ************

HHV7_RK_U4      TCAGCTCATTCAAGTTGACGTCTAACTCGCATGAGTTAAATGCCGTCCATGACAGATAAG 1556
A5_020313       TCAGCTCATTCAAGTTGACGTCTAACTCGCATGAGTTAAATGCCGTCCATGACAGATAAG 249
                ************************************************************

HHV7_RK_U4      GATGATTCGGTATAACATACGTAAT-GCGATTCCTTAGCGGCCCTTTACAAACATCATAA 1615
A5_020313       GGTGATTCGGTATANCATACGTTTTTGCGATTCCTTAGCGGCCCTTTACCNC-------- 301
                * ************ *******  * ************************

HHV7_RK_U4      TCGATTAAATCCAT 1629
A5_020313       --------------


HHV7U7

HHV7_RK_U7      AAAGTCCGGGTCCGGTTCTCGGTAATTTGGATGTGCATAACCAACAGCGCAAAGAAATTC 2460
A6_020424       -------TGGTCGGGTTCTCGGTAATTTGGATGTGCATAACCAACAGCGCAAAGAAATTC 53
                       **** ************************************************

HHV7_RK_U7      GTCTGCTCCTACCAAATTACTCCAATCATAAGTGATGACACTACGTCTACATTTGAAATG 2520
A6_020424       GTCTGCTCCTACCAAATTACTCCAATCATAAGTGATGACACTACGTCTACATTTGAAATG 113
                ************************************************************

HHV7_RK_U7      ATAAAATTCGCCAAAAATAAGCACATAGTTTGGCGGCCAAGCAAGACAGCACTTTATACC 2580
A6_020424       ATAAAATTCGCCAAAAATAAGCACATAGTTTGGCGGCCAAGCAAGACAGCACTTTATACC 173
                ************************************************************

HHV7_RK_U7      ATTGCGAACTATCTCCATGCCTTCTCCATATGAAGACAGCAGCAACTCAACGGCAAAATC 2640
A6_020424       ATTGCGAACTATCTCCATGCCTTCTCCATATGAAGACAGCAGCAACTCAACGGCAAAATC 233
```

```
                ************************************************************

HHV7_RK_U7      AAAAGACGGTAAATCCGGAGTCGGTTCAATGCAAACATCATT-GCACACAGGAACAATTT 2699
A6_020424       AAAAGACGGTAAATCCGGAGTCGGTTCAATGCAAACATCANTTGCACACAGGAACAA--- 290
                ****************************************** * **************
```

HHV7U8

```
HHV7_RK_U8      TTTCCTTCCAATTTGAATCGCTCCAATCAAGCACAAATTTTCCTTAAATCCTTGGTAATG 840
A7_020313       ----------------------CAATCAAGCNCAAAT--TCCTTAAATC--TGGTA-TG 32
                                     ********* *****  **********  ***** **

HHV7_RK_U8      GTCATTCAGTTCTGCCAATTGTTG-TTCTGAGTAACCAGCAACGTCAACTGGCTGTGCAA 899
A7_020313       GTCATNNN-TNCTGCCAATNGTTGGTNCTGAGTNNCCCCCN-CGTCA-CTGGCTGTGCA- 88
                *****    * ******** **** * ******  **  *  ***** ***********

HHV7_RK_U8      GCCTTAACCACAAATTCTTTGGCCAAATCAAAGACAAGCATTCATTTCGATGTCGTTCAG 959
A7_020313       GCCTTANCCACAAATTCTTTGGCCAA-TCCAAGACAGC--TTCATTTCGATGTCGTTC-G 144
                ****** ******************** ** ******    ****************** *

HHV7_RK_U8      TAAGAAGGCGAATTCCAGGGAAATCAGTAAAACGACATAAATTCTTCAATTCCCTTAACA 1019
A7_020313       TAG--AGGCGAATTCCNGGGAATCNTN----NCGAC-TATNTTC----ATTCCCTTA-CA 192
                **   ************ *****          **** **  ***    ********* **

HHV7_RK_U8      ATCCATCATTTAGTTGTCCAACACAAGGCAAAGTTTTATATTGT---TCTGCCATCGTCT 1076
A7_020313       TCNNNTNNNTTTTTTGTCCNCCCTGGGCAAGTNTTATATNGTNTGCCTCTCTCNTCNTNC 252
                     *   **  ******  *    *  *   ** ***  * *   ***  * ** *

HHV7_RK_U8      CAAAATTTTTCAT------------------------------ 1089
A7_020313       CCGCGGNNNNTGGCNGCNGGGGANCTGCACCTNGGCCNNTNC 294
                *
```

HHV7U10

```
HHV7_RK_U10     ATGGCTATAGCAGAAAGCGGAGCGTTCGAAATAAATGTCAATCTTGGAAAATCCATCGCT 60
A8_020502       ----ATTNAAATGAAAGCGNGGNANNGCTTCTTCATGTTATTTCTNG----TGTNCTCCT 52
                     *  *   *******  *          *  **** * *  * *    *       **

HHV7_RK_U10     TCCATACAAAAAAATCCTATAATCTACATGCGTAGACATCTCTCTTTTTATGTGGAGTTA 120
A8_020502       ACCANACNAATAA--NCTANTATANATCCGNTCNNTACNTTCTCTTTTATTGTGGAGTTN 110
                 *** ** ** **   ***  **  *   *          ********  *********

HHV7_RK_U10     TTAAAAT---TTATCATTCATCAATATGAGCAATGTTTTTTACCACCGAAAGGAACGATA 177
A8_020502       NTNAAATACCTNATCATTCATCAATATGAGCAATGTTTTTTACCACCGAAAGGAACGATA 170
                 * ****   * **************************************************

HHV7_RK_U10     CTTTACCACAATGGCTTAATCGAACTAAATACTTTAATTATCGATCTGAATCAACAAATT 237
A8_020502       CTTTACCACAATGGCTTAATCGAACTAAATACTTTAATTATCGATCTGAATCAACAAATT 230
                ************************************************************

HHV7_RK_U10     ACATCGAAACAGCAGATCTATAGTTGGACGAGCATTACTTTGCCAAAAATATTTTCGACT 297
A8_020502       ACATCGAAACAGCAGATCTATAGTTGGACGAGCNTTNCTTTGCC----NTNTTTTCGACT 286
                ********************************** ** *******     * *********

HHV7_RK_U10     AAAGAATTATATTTTATTGTCGCGTCACCAGAATCGGAAAACATAACTCTCAATCCAGCT 357
A8_020502       N---GATTATATTTTNTTGTCGCGTCNNN------------------------------ 312
                     ********** **********
```

HHV7U11

```
HHV7_RK_U11     TAAGATTGACGTAATCGGCTGCGTGTGTCACAGTGCCCCAACGAAGAAAACTGCCCAGTT 1860
A9_020424       ----------------------------------------------------TGCCCCAGTT 10
                                                                   *******

HHV7_RK_U11     TAAGCAAATCTTTTGCTACGTTTTTAAATTCTGTTTCAAAG-CCAACTTTGTTAAAAGAT 1919
A9_020424       TTAGCANATCTTTTGCTACGNTTTTTAATTCTGTTTCNAAGGCCAACTTTGTTNAAAGAT 70
                * **** ************** **** ************ *** ************ ******

HHV7_RK_U11     GTTAAACTAGAATAATCAGTAATCATCTGTCTTTGCTCTAAGTAATCTTTCAGAATTTTT 1979
A9_020424       GTTANACTAGAATAATCAGTAATCATCTGTCTTTGCTCTAAGTAATCTTTCAGAATTTTT 130
                **** ************************************************************

HHV7_RK_U11     TGCACCTTTAGTATTTCTTGCCAAACTTCCTCATAGTCTGGTTTCTTTTTACACAAGGTT 2039
A9_020424       TGCACCTTTAGTATTTCTTGCCAAACTTCCTCATAGTCTGGTTTCTTTTTACACAAGGTT 190
                ************************************************************

HHV7_RK_U11     TGATGAAAATAAATCCAAAGAATCAAATTGTTGTAAAGTGTTTTTACATTGTTAATGGAA 2099
A9_020424       TGATGAAAATAAATCCAAAGAATCAAATTGTTGTAAAGTGTTTTTACATTGTTAATGGAA 250
                ************************************************************

HHV7_RK_U11     TTTCCAGTTTTAACGATACATTGACTCAAAATCCATGGATTTTCGCGAATATCAACAACT 2159
A9_020424       TTTCCAGTTTTAACGATACATTGACTCAAAATCCATGGATTTTCGCGAATATCAACAACT 310
                ************************************************************

HHV7_RK_U11     GGTAAAGACGATATGTTTTCAAAAAATCGAGATAAAAAACACTTTGCTTCATCCGAAATC 2219
A9_020424       GGTAAAGACGATATGTTTTCAAAAAATCGAGATAAAAAACACTTTGCTTCATCCGAAATC 370
                ************************************************************

HHV7_RK_U11     CAAGCAAAGGGTAAATGTGACATTTTCAT 2248
A9_020424       CA--------------------------- 372
                **
```

HHV7U12

```
HHV7_RK_U12     CAGCATGATGAAGAGTACAAGTACAATTATACGTGTATTACGCCAACAGTACGGAAAGCC 180
A10_020413      ------------------------------------TTACGCCAACAGTACGGAAAGCC 23
                                                   ***********************
```

```
HHV7_RK_U12     CAAAGACTTGAAAGCGTAATTAACGGAATTATGCTAACGCTGATACTTCCTGTTAGTACT 240
A10_020413      CAAAGACTTGAAAGCGTAATTAACGGAATTATGCTAACGCTGATACTTCCTGTTAGTACT 83
                ************************************************************

HHV7_RK_U12     GTTGTCATATGCACTCTGCTAATCTACTACAAATGGACAAAACAGACAATTACTTCTCCA 300
A10_020413      GTTGTCATATGCACTCTGCTAATCTACTACAAATGGACAAAACAGACAATTACTTCTCCA 143
                ************************************************************

HHV7_RK_U12     TATCTTATCACACTCTTTATTAGTGATTCTTTACATTCATTGACTGTGTTACTTCTCACA 360
A10_020413      TATCTTATCACACTCTTTATTAGTGATTCTTTACATTCATTGACTGTGTTACTTCTCACA 203
                ************************************************************

HHV7_RK_U12     TTGAACCGAGAAGCTCTCACAAACCTTAATCAGGCTTTGTGTCAATGTGTGCTTTTTGTA 420
A10_020413      TTGAACCGAGAAGCTCTCACAAACCTTAATCAGGCTTTGTGTCAATGTGTGCTTTTTGTA 263
                ************************************************************

HHV7_RK_U12     TACAGTGCGTCCTGCACATACAGTCTGTGCATGCTAGCAGTAATATCCACAATACGCTAT 480
A10_020413      TACAGTGCGTCCTGCACATACAGA---------------------------------- 287
                ************************
```

HHV7U13

```
HHV7_RK_U13     ATGACTTTGAATCAACAGCCATCTAATTGCAGACTAATTACAGCTAACGATCCAGTACTT 60
A11_020419      ------------------TAAAATGGTGACGAACGAACTTCGTCGTCTGAGTCATAA--A 40
                                  *  *  *  *  ** ** * *  *     **  **  *

HHV7_RK_U13     GCATCGAATTTTAC--AATGCAGCCTACTTTTAAAATAGCCGATAAGAA------AGTTG 112
A11_020419      ACATGAAATCTGGCTTAATGGAATATGCTCTAATGCTATCGGATGTCTTCTTTTCAATTT 100
                 ***  *** *  *  **** *   * ** * *   ** * ***           * **

HHV7_RK_U13     TACTCAGAGATCATAATTACATTGCAGTCAAAGATTTTGTTCTATCAAGGTCTTTCA-TG 171
A11_020419      TTTCCTGACATC--GAATGCAATGCA-TGAAAGACCTTGATAGAACAAAATCTTTGACTG 157
                *   * ** ***   * * ** **** * *****  *** *  * ***  ***** * **

HHV7_RK_U13     CATTGCATTC--GATGTCAGGAAAAAATTGAAAAGAAGACATCCGATAGCATTAGAGCAT 229
A11_020419      CAATGTAATTATGATCTCTGAGTACAACTTTCT------TATCGGCTATTTTAAAAGTAG 211
                ** ** * *   *** ** *   * ** *          *** * **   * * ** *

HHV7_RK_U13     ATTCCATTAAGCCAGATTTCATGTTTTATGACTCAGACGACGAAGTTCGTTCGTCACCAT 289
A11_020419      GCTGCATT--GTAAAATTCGATGCA--AGTACTGGATCGTTAGCGGTAATTAGTCTGCAA 267
                  * ****  *  * ***  ***    *  ***    **     * *  ** ***  **

HHV7_RK_U13     TTTTATAA------ 297
A11_020419      TTAGATGGCTGTTG 281
                **  **
```

HHV7U14

```
HHV7_RK_U14     ATGGAACAGCAAAAAGGATTTTCGATTCCATTTTTTGTCACTGACGAAAACTGCAACTTT 60
A12_020308      ----------------------------------------TCGA--AAAACTGCAACNN- 17
                                                        **  ***********

HHV7_RK_U14     GTGCCTGAAATATTACCTCGTATACATACTAAA--TTTCTTAAAG-ATGTCTTAATTGCC 117
A12_020308      GTGCNNAAAANNATACCTCGTATACAAAANNAAAATTTCTTAAAGGATGTCTTAATTGCC 77
                ****   ***   ************** *   **  ********** ***************

HHV7_RK_U14     GATTCCTACAACTCTGTTAGTTGGGCAAACAGTTTTATTCCCATGCC-TATTCAA-ACGC 175
A12_020308      GATTCCTACAACTCTGTTAGTTGGGCAAACAGTNNGATTCCCATGCCGTATTCAAGACGC 137
                ***********************************   *********** ******* ****

HHV7_RK_U14     TTGAACAAA-TTATGGTTCTTATAACAAAGTTCAAGTTTTCTCGCTCGCGTGATTTTTTA 234
A12_020308      TGGAACAANATTATGGTTCTTATAACAAAGTTCAAGATTTCTCGCTCGCGTGATTTTTTA 197
                * ******  ************************** ************************

HHV7_RK_U14     TTCCCAGTAATTCGAT---TAGCTGTTCATATCA-ATAGGTTCCACACAGGGAAGAAACA 290
A12_020308      TTCCCAGTAATTCGANGGGNNGCTGTTCATATCNCATAGGTTCCACACAGGGAAGAAACA 257
                ***************      ************  *************************

HHV7_RK_U14     GCTGAAAACCA-TGATAGAAATTATGAAAAGCTTGTTTAACACCGAAGAG-GCTATG-CG 347
A12_020308      GCTGAAAACCNATGATAGAAATTATGAAAAGCTTGTTTAANNCCGTTGAGTNNNANGTCG 317
                **********  ******************************  ***  ***    * * **

HHV7_RK_U14     ACGATTCGATGAAGCATTGATGATTCTATTTTCTAATGAGCAAACCAATACTTACATGAC 407
A12_020308      ACGACCC---------------------------------------------------- 324
                ****  *
```

HHV7U15

```
HHV7_RK_U15     AAATTTTGAAACTCGTTTGTGTAGAAATAATAGAGGTGGACTCACCGTTTTACAGGTATG 420
B1_020419       ------------------------------------------------------------T 1

HHV7_RK_U15     TGTGTGAATTGAGC-GACGATATAACAATCCTCAGTTTTGTGGACATCTAGATTGAAAAT 479
B1_020419       TGTGTGAATTGAGCTNACGATATAACAATCCTCAGTTTTGTGAACATCTAGATTGAAAAT 61
                **************  ************************** *****************

HHV7_RK_U15     CTGCTTGATTTCTTGCCATCTCTTAATAATAAACAAAGCTTCAATCACTACAGACTTTGT 539
B1_020419       CTGCTTGATTTCTTGCCATCTCTTAATAATAAACAAAGCTTCAATCACTACAGACTTTGT 121
                ************************************************************

HHV7_RK_U15     CACAGTTGTGGTTAATGTTAATCCAGAAAAAATAAATCTATATGTTGTATTAAAATCCAT 599
B1_020419       CACAGTTGTGGTTAATGTTAATCCAGAAAAAATAAATCTATATGTTGTATTAAAATCCAT 181
                ************************************************************

HHV7_RK_U15     TTGAAAAAGATATTTTATATAGATTGTCTCCACTTTAGA-TATAATATTA-GATAAC-GT 656
B1_020419       TTGAAAAAGATATTTTATATAGATTGTCTCCACTTTAGNATATAATATTAAGATAACCGT 241
                ************************************  ********** ****** **

HHV7_RK_U15     CA--TCAATATCTGTAGTGGACACAGCTCGCGAAATTCCTGTAGTCGTTGTCTTCTCCAG 714
B1_020419       CCATCCAATATNC-TNGTAGNTGGGACCCC---AGCCTCGGCGAANAA------------ 285
```

```
                *    ******   * ** *      * *    *    * *

HHV7U17

HHV7_RK_U17     ACGCTCTTCTTCTCCGAGAAGAATGATGGTTTTAGCAGCTGCTGTTGTTACGTCGTTTTT 720
B2_020424       ---------------------------------TGCAGCTGCTGTTGTTACGTCGTTTTT 27
                                                  ***************************

HHV7_RK_U17     AGCACTCAGAAAACCGATGACTTGTAACCTTTCATTGCAACATAGGAGATAGCAACCTAT 780
B2_020424       AGCACTCAGAAAACCGATGACTTGTAACCTTTCATTGCAACATAGGAGATAGCAACCTAT 87
                ************************************************************

HHV7_RK_U17     GAAAACAACAAAAACTTATTAGAACATCATGACACATTGGAAATGTTTTGACTGCTTTTG 840
B2_020424       GAAAACAACAAAAACTTATTAGAACATCATGACACATTGGAAATGTTTTGACTGCTTTTG 147
                ************************************************************

HHV7_RK_U17     CAACTTACATAGGTCTTGCAGCTCGGTTTCACTAGTAGTTCTGAAAGGCATATCCCGCAA 900
B2_020424       CAACTTACATAGGTCTTGCAGCTCGGTTTCACTAGTAGTTCTGAAAGGCATATCCCGCAA 207
                ************************************************************

HHV7_RK_U17     AGGCGAAATGATTATTTTCCTACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATT 960
B2_020424       AGGCGAAATGATTATTTTCCTACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATT 267
                ************************************************************

HHV7_RK_U17     TTGAGTAACGAAGTCACTAAGAGAATCGCGATAAAAGGCTTGTTTAACAAGGCTTAAATT 1020
B2_020424       TTGAGTAACGAAGTCACTAAGAGAATCGCGATAAAAGGCTTGTTTAACAAGGCTTAAATT 327
                ************************************************************

HHV7_RK_U17     TGCGAGTGCATGCCTAAATCGAACTTTTTGTTCACTTTCGGCCAT 1065
B2_020424       TGCGAGTGCATGCCTA----------------------------- 343
                ****************

HHV7U18

HHV7_RK_U18     AGTGATGGAACCACCCGAATTATGACCTTCTAACCAACATTCAAGTGTACTGTTAGCCCG 540
B3_020424       ----------------------------------------------------TTGTTAGCCCG 11
                                                                     **********

HHV7_RK_U18     AGAAGCAAAGTGCCATTTTTTGGCAATTAGTGCTCTTAACATAACTCCTCTAAAGTAATA 600
B3_020424       AGAAGCAAAGTGCCATTTTTTGGCAATTAGTGCTCTTAACATAACTTCTCTAAAGTAATA 71
                ********************************************* *************

HHV7_RK_U18     ATTGATCAGCACTCTCAAATAGGTTAATTTATGAGATGTTGTAGTAAGATTTACTAGACT 660
B3_020424       ATTGATCAGCACTCTCAAATAGGTTAATTTATGAGATGTTGTAGTAAGATTTACTAGACT 131
                ************************************************************

HHV7_RK_U18     CATTTCGCGAAATCCTTCAGTTTTTAAAGACAAAGAAATCAGAAAAGCTGTTCTTACAGA 720
B3_020424       CATTTCGCGAAATCCTTCAGTTTTTAAAGACAAAGAAATCAGAAAAGCTGTTCTTACAGA 191
                ************************************************************

HHV7_RK_U18     TAAACAACTAGAATTCCCTATTAGCATTTCATTAACATAACACGTGGTGACACAAACAGC 780
B3_020424       TAAACAACTAGAATTCCCTATTAGCATTTCATTAACATAACACGTGGTGACACAAACAGC 251
                ************************************************************

HHV7_RK_U18     TTCTAATTTATTATCTAGAAATATTATATTCTGGTGACATTTCAAATTTGACATGAGTGT 840
B3_020424       TTCTAATTTATTATCTAGAAATATTATATTCTGGTGACATTTCAAATTTGA--------- 302
                ************************************************

HHV7U19

HHV7_RK_U19     CTGAGACAAGGACAGTAAAAAAACTATTTGTCTTCGCGGCGTCCCAGAAGCTGACACCCA 720
B4_020413       ----------------------------------------------TAAGCTGACACCCA 14
                                                               **************

HHV7_RK_U19     TT-CTTCGGCAACCCCAAACATCACCATCTTTCCCTCAAATGGATAAAATCCGATGTTAA 779
B4_020413       TTTCTTCGGCAACCCCAAACATCACCATCTTTCCCTCAAATGGATAAAATCCGATGTTAA 74
                ** ********************************************************

HHV7_RK_U19     CTAAACGTTTAAGTCTCCTTTCTGGAAGAATGTCAGTCATATCACAAATTCGAATATTGA 839
B4_020413       CTAAACGTTTAAGTCTCCTTTCTGGAAGAATGTCAGTCATATCACAAATTCGAATATTGA 134
                ************************************************************

HHV7_RK_U19     TACACATTATGTGTTGAATGGACACGCCATCCTTCATTATTTCTTGAACTGTTTTTTCAA 899
B4_020413       TACACATTATGTGTTGAATGGACACGCCATCCTTCCATTTT------------------- 175
                ************************************  * **

HHV7U20

HHV7_RK_U20     AATGAATTCAGCGTTAAAATTG--TGGTGATAATGAACACTGCGATTCCTATGTGACCAG 898
B5_020502       ------TTCNNCGTTAAAATTNGTNGGNGATAATGAACACTGCGATTCCTATGTGACCAG 54
                      ***  **********    ** *********************************

HHV7_RK_U20     ACAAAATAAGAGAAAATGCTGTTTAATGTAGATTCGTCCTTGTTGGTAGAAAAGACGTAA 958
B5_020502       ACAAAATAAGAGAAAATGCTGTTTAATGTAGATTCGTCCTTGTTGGTAGAAAAGACGTAA 114
                ************************************************************

HHV7_RK_U20     GACAAAGCATTAAATTTCGTTGGAAAATCTGTGAGGAAAGACGAAATACTAGTATTGCCG 1018
B5_020502       GACAAAGCATTAAATTTCGTTGGAAAATCTGTGAGGAAAGACGAAATACTAGTATTGCCG 174
                ************************************************************

HHV7_RK_U20     ACAGACAGATTCTTATAATGTGCCAGTTCATAGTTATCCATCATCAATTCATATTTCTCA 1078
B5_020502       ACAGACAGATTCTTATAATGTGCCAGTTCATAGTTATCCATCATCAATTCATATTTCTCA 234
                ************************************************************

HHV7_RK_U20     AGATGTTCTTTATTGATTATTGACATTGTCA-ACTTATGTGCCACATATTTGGTAAAATG 1137
B5_020502       AGATGTTCTTTATTGATTATTGACATTGTCATACTTATGTGCCACAA------------- 281
                ********************************** ***************
```

```
HHV7U21

HHV7_RK_U21     GGGTGAACAGATAATTTGAATATGAATGTAAGACTCGACATCAAGACAGAGTTTATCATA 960
B6_020413       ------------------------------------------------------------T 1


HHV7_RK_U21     TCGCT-TCATCAGTCCTTTTGTTTTTTTAAAGTTCTTAAACAAAAAATGATTTAGTGTAG 1019
B6_020413       TCGCTGTCATCAGTCCTTTTGTTTTTTTAAAGTTCTTAAACAAAAAATGATTTAGTGTAG 61
                ***** ************************************************************


HHV7_RK_U21     AAAATACTTGATAACAGTTGAAAGCCAAATATTTGCCAAAATCGGTTTCAATTTGCGTTT 1079
B6_020413       AAAATACTTGATAACAGTTGAAAGCCAAATATTTGCCAAAATCGGTTTCAATTTGCGTTT 121
                ************************************************************


HHV7_RK_U21     CGCTGTTGCCAAAAAAGAAAGCATTGTCTGTCATTGCCGCGACTGCAGAGTGTCTGATTT 1139
B6_020413       CGCTGTTGCCAAAAAAGAAAGCATTGTCTGTCATTGCCGCGACTGCAGAGTGTCTGATTT 181
                ************************************************************


HHV7_RK_U21     TATCGTCTGAACATTGTTTGCTCAAAGATATGTCGAACAATAGCAAATCATCCACTGTAG 1199
B6_020413       TATCGTCTGAACATTGTTTGCTCAAAGATATGTCGAACAATAGCAAATCATCCACTGTAG 241
                ************************************************************


HHV7_RK_U21     CATTCACATCAAAATCCACAACTGCTAAGGGACAGAAATCCGGATAGAGTTCTCCATAAA 1259
B6_020413       CATTCACATCAAAATCCACAACTGCTAAGGGACAGAAATCCGGATAGAGTTCTCCATAAA 301
                ************************************************************


HHV7_RK_U21     TCACAGGAACGCAAAACAGCAAGATAGTCCACAT 1293
B6_020413       TCACAGGAACGCAAAACAGCAAGATA-------- 327
                ***************************


HHV7U23

HHV7_RK_U23     ACTAACTGCTGCAATTAAAATCACAAGGAACAAAACGAACTTTTTTAAAGGGCTATTCAA 180
B7_020413       -----TTGCTGCAATTAAAATCACAAGGAACAAAACGAACTTTTTTAAAGGGCTATTCAA 55
                     ************************************************************


HHV7_RK_U23     AAGTTGAGGTTTTGTCAGTTCTGGAGTTGAAATTGTCAGTGTGAAAAATTCAGATGCATT 240
B7_020413       AAGTTGAGGTTTTGTCAGTTCTGGAGTTGAAATTGTCAGTGTGAAAAATTCAGATGCATT 115
                ************************************************************


HHV7_RK_U23     CCCACTGTTTTCATGATTAGTAGTACTGCCACTGGAGAATGTCGAGGTCTTTGTAAAATT 300
B7_020413       CCCACTGTTTTCATGATTAGTAGTACTGCCACTGGAGAATGTCGAGGTCTTTGTAAAATT 175
                ************************************************************


HHV7_RK_U23     TAAGGATGTAGATTCGACATTCTCTTTAGTTGTTGCGATAGTCGTAGGTATACTCATAGA 360
B7_020413       TAAGGATGTAGATTCGACATTCTCTTTAGTTGTTGCGATAGTCGTAGGTATACTCATAGA 235
                ************************************************************


HHV7_RK_U23     CACAGACATTAAATCGCTCGTACTCTCTGTATCTGATGTGATTGACGTGGAACTAAAATC 420
B7_020413       CACAGACATTAAATCGCTCGTACTCTCTGTATCTGATGTGATTGACGTGGAACTAAAATC 295
                ************************************************************


HHV7_RK_U23     TGAAGTTCTTGACATTTCCAGAGTGGGCGTACTAGAAACAGTACTCATGTTAACAAAGGT 480
B7_020413       TGAAGTTCTTGACATTTCCAGAGTGGGCGTACA---------------------------- 328
                *********************************


HHV7U24

HHV7_RK_U24     TCAGGGCTTCGTACCCCCATATCGAACATTGAAGATGTGTAGAATAAGAATCAGACATAA 60
B8_020424       ------TTTCGTACCCCCATATCGAACATTGAAGATGTGTAGAATAAGAATCAGACATAA 54
                      ************************************************************


HHV7_RK_U24     GATAACCGATATTATGAGACATGCCAGTATTATAAAAGTCCCTCTTCGGCGCACATTTTT 120
B8_020424       GATAACCGATATTATGAGACATGCCAGTATTATAAAAGTCCCTCTTCGGCGCACATTTTT 114
                ************************************************************


HHV7_RK_U24     GATTTCAGAACTAGAAGTTGAATTTATTGTTCTTGGTGATAGATTTTCCTGATGCAAAA 179
B8_020424       GATTTCAGAACTAGAAGTTGAATTTATTGTTCTTGGTGATAGATTTTCCTGATGCAA-- 171
                ************************************************************


HHV7U24A

HHV7_RK_U24A    CTAGGGATATTTAGAGCACCTGATAAATGATAACATATTCTCTG-CTATTG--AAGGACT 57
B9_020314       -TAGGGATATTTAGAGCACCTGATAAATGATAACATATTCTCTGTCTATTGTNNAGGACT 59
                 ********************************************* ******   ******


HHV7_RK_U24A    TATATA-ATCATGATATGCAATGCAAGCCAAAATAAACAGAACAAACAAATAATAAATCC 116
B9_020314       TATATCCATCATGATATGCAATGCAAGCCAAAATAAACAGAACAAACAAATAATNA-TCC 118
                *****  ************************************************ * ***


HHV7_RK_U24A    CTTTCCAACAATCTTCATAGGTCTCTGGGAAAAAATGGTTGTACTTATAAAACAT----- 171
B9_020314       CTTTCCTACAATCTTCATT--TTTTTTTTNANTCTTTTTTTTNNNCATNATTTATCCNTC 176
                ****** ***********   * * *    *    *  ** *    ** *   **


HHV7_RK_U24A    ---------------------------------------------
B9_020314       CTCGTGCNNCTTGNNNTTTTTTCTNATTTTNTGTTGTCCNACA 219


HHV7U25

HHV7_RK_U25     AAAACCATTGTAGACATACAAATGCCCTCTGTCACCTTGTAAAAGAATTGGTCCTCTGCA 720
B10_020419      ------------------------------------------------------------CN 2
                                                                             *


HHV7_RK_U25     AGTTGATTTTTTAACATG-ATGTAACACAC-----CAACAACGATCAACGTTTTTTCACA 774
B10_020419      GATNGGATTTTTAACACNTAGGTGATANACACCCNCANCAACGATCAACGTTTTTTCACA 62
                  * *  *********   * ** * * **     ** ************************
```

```
HHV7_RK_U25     ACACAAGTAAGTGTTTCGAATTTCTTCGATGTCTTCCTTTGAAATTTCTGGGATTGTTTC 834
B10_020419      ACACAAGTAAGTGTTTCGAATTTCTTCGATGTCTTCCTTTGAAATTTCTGGGATTGTTTC 122
                ************************************************************

HHV7_RK_U25     TAATGAGTGAAATGTAACTCGCCATTCTTCTGGGAAAGGCAATACGAAAGATGTACCTTC 894
B10_020419      TAATGAGTGAAATGTAACTCGCCATTCTTCTGGGAAAGGCAATACGAAAGATGTACCTTC 182
                ************************************************************

HHV7_RK_U25     CTTACTTAAGATGAATTCTTCGAGAAGAGTAAGTTTATTGACACAGCACAAAT--TCGCT 952
B10_020419      CTTACTTA-GATGA-TTCTTCGAGA-GAGTACGTNTCNCGCCNCAGCNCNNNNCTTCGCT 239
                ******** ***** ********** ***** ** *   * * **** *      *****

HHV7_RK_U25     AA---GAAATCCAT---------------------------------------------- 963
B10_020419      NCATCNNAATTCCCGCNGGCCNGCCNNTGGGCGGGCCCGGGGNGCNNTGCGGACGTCGNG 299
                   *** *

HHV7U26

HHV7_RK_U26     ATAAGATGCCATTGAATGCGTTG-GAAATATTAGCATGCATCTACAGACTATGGCTCTCA 599
B11_020424      -------TCCATTGAATGCGTTGTGAAATATTAGCATGCATCTACAGACTATGGCTCTCA 53
                       *************** ************************************

HHV7_RK_U26     AGCGGATTTCAGAAAAAAATAAAAACGAGCAGATTACCAAAAAAAATTCATCGGTCATTT 659
B11_020424      AGCGGATTTCAGAAAAAAATAAAAACGAGCAGATTACCAAAAAAAATTCATCGGTCATTT 113
                ************************************************************

HHV7_RK_U26     GCAACACGAGTAATGATAAAGTTATAATTTTGAGATCGTTGAAAGGACAATAGTTAAACA 719
B11_020424      GCAACACGAGTAATGATAAAGTTATAATTTTGAGATCGTTGAAAGGACAATAGTTAAACA 173
                ************************************************************

HHV7_RK_U26     AATTCACCCGACGGTAAGTCAATAGCAAACCTCCGCTAAAAAGGAATATTAGCACTACTT 779
B11_020424      AATTCACCCGACGGTAAGTCAATAGCAAACCTCCGCTAAAAAGGAATATTAGCACTACTT 233
                ************************************************************

HHV7_RK_U26     GTTCAAACAGCGTTTTCTTTATCATAAGATGGAA-ACAGCTTTCTCCCGGAATTACATTC 838
B11_020424      GTTCAAACAGCGTTTTCTTTATCATAAGATGGATTACAGCTTTCTCCCGGAATTACATTC 293
                ********************************  *************************

HHV7_RK_U26     CCTAAAACAAGGCCCATAGAAAAACTTATCATTACATTCCACAT 882
B11_020424      CCTAAAACNNGGCCCA---------------------------- 309
                ********  ******

HHV7U27

HHV7_RK_U27     CATAAAACTATCAAATAATGGAACAAAGTTATTTATCGTTTTAGTCGAAAAATGATCTGT 840
B12_020424      ------------------------------------------------------TGTAT 5
                                                                      * * *

HHV7_RK_U27     ATCGGTTATATATAAACATTCTG-CATGAATAATCAG-CTTTAAAACCAAATGGTTTTTT 898
B12_020424      TCGGNTTATATATAAACATTCTGGCATGNATAATCAGGCTTTAAAACCAAATGGTTTTTT 65
                   * ****************** **** ******** **********************

HHV7_RK_U27     GCAGATTGAATTATGATTGCCGGTTGAGGAGTAAAAGTAACTGTAGTATTTTCTCGCAAA 958
B12_020424      GCAGATTGAATTATGATTGCCGGTTGAGGAGTAGAAGTAACTGTAGTATTTTCTCGCAAA 125
                ********************************** *************************

HHV7_RK_U27     AGCTTTGCAAAAGCTCTTAGAGGCTTATTGATTGTTTTCCAAGTTTTCATGTGAAA--AG 1016
B12_020424      AGCTTTGCAAAAGCTCTTAGAGGCTTATTGATTGTTTTCCAAGTTTTCATGTGNAATAAG 185
                ********************************************************** **  **

HHV7_RK_U27     CCATG---GTGGGAGGCTCTTTGGAATCTCGATGTTCGCGA-TGATC-TCGGTGAC-TAT 1070
B12_020424      CCGAGTGGNTGGGAGGCTNNTTGGAATCTCGATGTTCGCGNATGATCGTCGGTNACGTAT 245
                **  *    **********  **********************  ***** ***** ** ***

HHV7_RK_U27     GGTGTTCTCTGTTGCTACGATCCAT- 1095
B12_020424      GGNGTATCGTCGTNNTGTCGTACGAA 271
                ** **    *  *  *    * *

HHV7U28

HHV7_RK_U28     ACTTTCTAAAATAAGTGAGATGGAATCAAATTCCGAAGCAGAGTTTGTTTTGAGATAGAA 2160
C1_020228       -------------CGNNGNGAGNAATTNA--TGGGCATCANCGTAAGCACCAAAAANNAG 45
                           *      * ***  *  *  * * **  **  *     * *   *

HHV7_RK_U28     CTGGCCTAACAAACGATCAACGTGCGGTGCATCATGCATTGCTCCCAGTGCTGAAATTAT 2220
C1_020228       CCGGTTTATTTTNTT-TGTTTGTCTNGTNNAN-GTATATT-TTTTGTGTGGCGAGAG-AT 101
                * **  **         *    **   **  *   *  ***  *    ***  ** *  **

HHV7_RK_U28     GGCATCACGTAAGCACCACATATCTAAGCCCGAATTTAAGCGCCGGAGAATTTTTGAAAT 2280
C1_020228       GGTGNGGCG-ATTTANTGTAGATCTTGGNTCNTNTT-----ANTANATANGT----ANAT 151
                **     ** *   *    * ****  *  *   **          * *  *    * **

HHV7_RK_U28     TTCATATAACTGCTTATAAAGAAGAGTTTTGTGATGGAAGCGTGAATAGCAAAACCGTTT 2340
C1_020228       GTAGTTTNTTTGTTTGTATGGN----TTTNGTCN-----GTGTANATAGCAAAACCGTTT 202
                 *  * *   ** ** **  *     *** **       * **  ***************

HHV7_RK_U28     TTTGTTATATTCATCATGACAAAGACATTTCCTCAAGAGACCCATGTTTTTATTTATGGA 2400
C1_020228       TTTGTTATATTCATCATGACTA-GACATTTCCTNCNAGAANNCCCNNTTTTA-------- 253
                ******************** * **********      *  *    *****

HHV7U29

HHV7_RK_U29     ACCTTGCATTACACAGGATATTAGACAATTAGTCCCTTGAGAGATATAATTAGTTACTCC 480
C2_020424       ------------------------------------------------------TCTCC 5
                                                                        ****
```

```
HHV7_RK_U29     AAACAAAGACACCTTTTCTTCTAAGTAAGAACAACTATCTAACATTGTGTTGATACTGTC 540
C2_020424       AAACAAAGACACCTTTTCTTCTAAGTAAGAACAACTATCTAACATTGTGTTGATACTGTC 65
                ************************************************************

HHV7_RK_U29     AACTAATGAAGTTTTATACAGATTAGCTAAAGAAGTCCTCGTGCAGAAAGACACATTGAA 600
C2_020424       AACTAATGAAGTTTTATACAGATTAGCTAAAGAAGTCCTCGTGCAGAAAGACACATTGAA 125
                ************************************************************

HHV7_RK_U29     GGCATTCATGGCTAAAAACAATACACTAATCAACAAAAATTTTGGACTCTCCGCGTCTGT 660
C2_020424       GGCATTCATGGCTAAAAACAATACACTAATCAACAAAAATTTTGGACTCTCCGCGTCTGT 185
                ************************************************************

HHV7_RK_U29     TTCAAGTCTGAACATAAGCATTCCTAAATGGTCTCGTTCAGGTCCTAAAAGTTCGCAAGG 720
C2_020424       TTCAAGTCTGAACATAAGCATTCCTAAATGGTCTCGTTCAGGTCCTAAAAGTTCGCAAGG 245
                ************************************************************

HHV7_RK_U29     AGTAAAACTTCGCGAAAGCACACACTCTTTGGGAACACCAAAAGCTAGCAAACTTCTGAT 780
C2_020424       AGTAAAACTTCGCGAAAGCACACACTCTTTGGGAACACCAAAAGCTAGCAAAA------- 298
                ************************************************

HHV7U30

HHV7_RK_U30     GATAATACATTCTTAACTATAACAGCGAAAATTGAAATTTTTTCTCTTGCGGCTGAAAGT 120
C3_020413       ----------------------------------------------TTGCGGGTGAA-GT 13
                                                              ****** **** **

HHV7_RK_U30     ATCACAGCAGATAAGATATTGTCATTTGCGAAATTATTACCAATACACAATTATCATTTC 180
C3_020413       ATCACAGCAGATAAGATATTGTCATTTGCGANATTATTACCAATACACAATTATCATTTC 73
                ********************************* ****************************

HHV7_RK_U30     AACTTTATTAAAAATCATATGGTGTTTTATATATTAAATTATAGTACTCTAGGATTTGCC 240
C3_020413       AACTTTATTAAAAATCATATGGTGTTTTATATATTAAATTATAGTACTCTAGGATTTGCC 133
                ************************************************************

HHV7_RK_U30     AAAAGATTTTCAATAGCTGAAACACTATGTAGAGAATTAAAAATTTTCCACAGAACATTC 300
C3_020413       AAAAGATTTTCAATAGCTGAAACACTATGTAGAGAATTAAAAATTTTCCACAGAACATTC 193
                ************************************************************

HHV7_RK_U30     CAATCAACGTCTAAATTACAAAACTTAAATAATGCAGAAGTCCTTTATCAATTTGAAAAA 360
C3_020413       CAATCAACGTCTAAATTACAAAACTTAAATAATGCAGAAGTCCTTTATCAATTTGAAAAA 253
                ************************************************************

HHV7_RK_U30     CTTATTGAATCCATTCAAGTATTCAAGGCGGAATTGGCATCGCCTATTGGTAAACTTCTA 420
C3_020413       CTTATTGAATCCATTCAAGTATTCA-GGCGGAATTGGCATCGCCTATTA---------- 301
                *************************** ************************

HHV7U31

HHV7_RK_U31     ATGAGAATCATAGCAGGAAGTACAAATCAAAACGATCCTAAATACGGACCAAGAGCCGGA 60
C4_020502       -----------TGCAGGAAGTACAAATCAAAACGATCCTAAATACGGACCAAGAGCCGGA 49
                           **************************************************

HHV7_RK_U31     AAGCAATGTATGTCAAATTGTTTTTCTTTCTTGCATACAGTTTATTTGAACGGAATAAAC 120
C4_020502       AAGCAATGTATGTCAAATTGTTTTTCTTTCTTGCATACAGTTTATTTGAACGGAATAAAC 109
                ************************************************************

HHV7_RK_U31     AATGTGTTAAATAAAGAGTCTATTGACATAATCATGGAAAATGGAGCATTATTGGATAAT 180
C4_020502       AATGTGTTAAATAAAGAGTCTATTGACATAATCATGGAAAATGGAGCATTATTGGATAAT 169
                ************************************************************

HHV7_RK_U31     ATCAGTACAACGACATTGAAACTCGAAACTGGCAATATCCCAGAATATCGATTTTTCACA 240
C4_020502       ATCAGTACAACGACATTGAAACTCGAAACTGGCAATATCCCAGAATATCGATTTTTCACA 229
                ************************************************************

HHV7_RK_U31     GAAATCCCAAAAAAAATTAGTTCTAAT-TTTGGCGAAACAATACATGAATTATCTAGACC 299
C4_020502       GAAATCCCAAAAAAAATTAGTTCTAANATTTGGCGAAACAAA------------------ 271
                ***************************  **************

HHV7U32

HHV7_RK_U32     CAAAAGTTAGAATTTTTGGAAGTTGATGTTTTTCTTGTTTGTTCAAAACAATCGGCACCA 960
C5_020419       ------------------------------------------------------------T 1

HHV7_RK_U32     AGTTATTATTTTCTAGCGGGATTAAAAAACGATCGAAACTGTAACAATAATTTTCTATCA 1020
C5_020419       AGTTATTATTTTCTNGCGGGATTAAAAAACGATCGAAACTGTAACAATAATTTTCTATCA 61
                ************** ************************************************

HHV7_RK_U32     TGGCCATGGCCATAGCAAAGATAAAATTTCCACATGGAACTGTTAAAGAACACCCATGAA 1080
C5_020419       TGGCCATGGCCATAGCAAAGATAAAATTTCCACATGGAACTGTTAAAGAACACCATGAA 121
                ************************************************** *******

HHV7_RK_U32     ATATATCTTGATATGTCAGGTTGCTTGCTCTATGAGCGTATCGCAGTGTGATGTCATTAC 1140
C5_020419       ATATATCTTGATATGTCAGGTTGCTTGCTCTATGAGCGTATCGCAGTGTGATGTCAA--- 178
                ************************************************************

HHV7U33

HHV7_RK_U33     TCCACAATTTGTTTCCAAGTAAGACGTTGAATTTTGTAAAGAGCCAGAATTTTTAGAACC 120
C6_020502       ----------------------------------------------------TNNNAANNN 9
                                                                   *    *

HHV7_RK_U33     TTGCATTGCCTGTACAAAAAATACAAATATGCCAAAAGCAAAAAGAAATTTTTTGTCATC 180
C6_020502       ATGCATTGCCTGTACAAAAAATACAAATATGCCAAAAGCAAAAAGAAATTTTTTGTCATC 69
                 ************************************************************

HHV7_RK_U33     ATCATTTTATTAAAAATAGCTTTAATCTGATTTCTTGAATCCAGATCACACTTTAACCAA 240
```

```
C6_020502       ATCATTTTATTAAAAATAGCTTTAATCTGATTTCTTGAATCCAGATCACACTTTAACCAA 129
                ************************************************************

HHV7_RK_U33     ACTTTTAAAAACTTCCGGTAGAAATTATCAG-TTAAATCTTTACTCAAAAACTTTTTATT 299
C6_020502       ACTTTTAAAAACTTCCGGTAGAAATTATCNGGTTAAATCTTTACTCAAAAACTTTTTATT 189
                ****************************** * **************************

HHV7_RK_U33     CGTTACGAACATCGATAATACGCAATTTTCAGATT-TGCTGACATCTCCAATAAGAAATA 358
C6_020502       CGTTACGAACATCGATAATACGCAATTTTCAGATTGCGACANNCCATAGAATACGCAAAA 249
                ************************************  *         *  **** * ** *

HHV7_RK_U33     ACTGCAAATGACTGTACGAACCATGTTTACACAGCGGAGAAATCAGCCGACAAAGCAAAT 418
C6_020502       CGT---------------------------------------------------------- 252
                  *

HHV7U34

HHV7_RK_U34     AAGAGCGATATACTCAGATTCACTCTCGGGACCAAAAGCCACAAATTCAACTATAAAAGC 360
C7_020419       ------------------------------------------------------------T 1

HHV7_RK_U34     CGCATCTCGGTTTC-GTCTAATAATGCACTTTGTCAGTACACCACTTGAAGAAGGTTTAA 419
C7_020419       CGCATCTCGGTTTCTGTCTAATAATGCACTTTGTCAGTACACCACTTGAAGAAGGTTTAA 61
                ************** ************************************************

HHV7_RK_U34     TAGCTTCAATGTCATTCAAACACACAGATAAACACTCATTGACATTTAAGATGTTGCATT 479
C7_020419       TAGCTTCAATGTCATTCAAACACACAGATAAACACTCATTGACATTTAAGATGTTGCATT 121
                ************************************************************

HHV7_RK_U34     CTTTGACAATGTTCTGACTTTCTTGTTTTATGGGAGCGTGGAGATAACAAGACATGGAAA 539
C7_020419       CTTTGACAATGTTCTGACTTTCTTGTTTTATGGGAGCGTGGAGATAACAAGACATGGAAA 181
                ************************************************************

HHV7_RK_U34     CAGCACCTCCCGTATTTTTAAAAATAAATCTCGCTTCTTGTTTTGTCAAATCTTCCCAGA 599
C7_020419       CAGCACCTCCCGTATTTTTAAAAATAAATCTCGCTTCTTGTTTTGTCAAATCTTCCCAGA 241
                ************************************************************

HHV7_RK_U34     GACTCAGCAAATATTCCAGGCTATATGCAAAATCCAATTTTAAAAGTATATCACATAAAG 659
C7_020419       GACTCAGCAAATATTCCAGGCTATATGCAAAATCCAATTTTAAAAGTATATCACATAAAG 301
                ************************************************************

HHV7_RK_U34     GAAATTTAGGATTTTTAGAAAATAAATTTTTATCTGTGACACGAAAGTCTGCGGGACCCA 719
C7_020419       GAAATTTAGGATTTTTAGAAAATAAATTTTTATCTGTGACACGAAAGTCTGCGGGAA--- 358
                ************************************************************

HHV7U35

HHV7_RK_U35     AATTGAAGTGAGTAGTTTGATTTTTTTTT-TGATAATCTC-CGAGTGACTGCA-TTGATT 117
C8_020419       ---CGCACGGCGGCGCNCANTTTATTCACGTGATTATNTCGCCAGTGATCGTACTTGATT 57
                    * *  * *  *     *** **    **** ** ** * *****  * * ******

HHV7_RK_U35     TG-CTCTGTGATGTAAATATTTTAAATAGGTTAAGTGTCTGTATCCAACATTAATAAAAT 176
C8_020419       TGGCTNCGTGATGTAAATATTTTAAATAGGTTAAGTGTCTGTATCCAACATTAATAAAAT 117
                ** **  ************************************************************

HHV7_RK_U35     GTAGTCGGGCATCTAATGAAGGCAGCATTATATCGTACAGTTTTGGTAACATAGGGAGAA 236
C8_020419       GTAGTCGGGCATCTAATGAAGGCAGCATTATATCGTACAGTTTTGGTAACATAGGGAGAA 177
                ************************************************************

HHV7_RK_U35     AAGCTATTTCAATTTCAACTCGGGAAGATAATTGTTCATAAAGAATTTCTATATCTGTTT 296
C8_020419       AAGCTATTTCAATTTCAACTCGGGAAGATAATTGTTCATAAAGAATTTCTATATCTGTTT 237
                ************************************************************

HHV7_RK_U35     TGGCATTGTCGTTAGACAT 315
C8_020419       TGGCATTGTCGTTA----- 251
                **************

HHV7U36

HHV7_RK_U36     ----------------ATGGCTTATAAGGGAT--GGAACTCCGATTC-CTTTTCTATGAA 41
C9_020502       CCACCGCAGGCNGCCGCGAATTCACTAGTGATTTGGAACTCCGATNNTCTTTTCTATGAA 60
                                * *  ** ***  ************   ************

HHV7_RK_U36     CTCAGAACTTTTTAATGAAATTCTTTTATATGCACACTTAGATTCAAGTGGAATAGATTC 101
C9_020502       CTCAGAACTTTTTAATGAAATTCTTTTATATGCACACTTAGATTCAAGTGGAATAGATTC 120
                ************************************************************

HHV7_RK_U36     TGATGATCTTAACACAAATCCCAATACACTTGAAAATGAAATTAACTCCGTTGAAAAAAC 161
C9_020502       TGATGATCTTAACACAAATCCCAATACACTTGAAAATGAAATTAACTCCGTTGAAAAAAC 180
                ************************************************************

HHV7_RK_U36     TTTAAATATTGAAGAACTGAAAAAAATTACAACTGCTTTAAATATTGACAACCGATGCAA 221
C9_020502       TTTAAATATTGAAGAACTGAAAAAAATTACAACTGCTTTAAATATTGACAACCGATGCAA 240
                ************************************************************

HHV7_RK_U36     TATATGTTCAATTATAAACATCTGTTTACGACATGAAACCGATAAAATGTGGATCTATGA 281
C9_020502       TATATGTTCAATTATAAACATCTGTTTACGACATGAAACCGATAAAATGTGGATCTATGA 300
                ************************************************************

HHV7_RK_U36     CTATGCTCTTTTGTGTTACAAATGTAATGCTGCACCTAGAACTCCCTTGGCTGTCGTAAT 341
C9_020502       CTATGCTCTTTTGTGTTACAAATGTAATGCTGCACCTAGAACTCCCCCA----------- 349
                ********************************************

HHV7U37

HHV7_RK_U37     ATGATTATTCAAAGTACACGTAGACTGAGACGAGCATCTAGCTTGTTAAAGAAAAGCAAA 60
C10_020424      -------------TACACGTAGACTGAGACGAGCATCTAGCTTGTTAAAGAAAAGCAAA 46
                            ********************************************
```

```
HHV7_RK_U37     CCTTACAACAAAGAAAAAACTAACTTATCTTTATCTTTGTCACTTAAAGAACTCCATTCG 120
C10_020424      CCTTACAACAAAGAAAAAACTAACTTATCTTTATCTTTGTCACTTAAAGAACTCCATTCG 106
                ************************************************************

HHV7_RK_U37     GTTTTCAAATTATTTCCAGAGTATGAATTGAAATTTCTAAATATGATGAAACTTCCAATA 180
C10_020424      GTTTTCAAATTATTTCCAGAGTATGAATTGAAATTTCTAAATATGATGAAACTTCCAATA 166
                ************************************************************

HHV7_RK_U37     ACCGGTAAAGAACCTATCAAAATTCCATTCGATCTAAGTCTACATCATCAACATACGTGC 240
C10_020424      ACCGGTAAAGAACCTATCAAAATTCCATTCGATCTAAGTCTACATCATCAACATACGTGC 226
                ************************************************************

HHV7_RK_U37     TTAGACTTATCACCATATGCCAATGAACAAGTTTCAAAAAGTGCATGTGTTAATTGTGGT 300
C10_020424      TTAGACTTATCACCATATGCCAATGAACAAGTTTCAAAAAGTGCATGTGTTAATTGTGGT 286
                ************************************************************

HHV7_RK_U37     ACAACAAACATTCCAACAGCTTCAGATGCTATGGTGGCATATATGAATCAAATTTCAAAC 360
C10_020424      ACAACAAACATTCCAACAGCTTCAGATGCTATGGTGGCAA-------------------- 326
                ****************************************

HHV7U38

HHV7_RK_U38     ATTACCAGATGGAATAACGAAATGTCTATATTGGAAAGGGATGTTTTCTATAGAATCAGT 2760
C11_020424      --TACCAGATGGAATAACGAAATGTCTATATTGGAAAGGGATGTTTTCTATAGAATCAGT 58
                  ************************************************************

HHV7_RK_U38     AAATAATAAAGATTCGGATGCATCATAGATATGAAATTTTAAAGGCGCACGCAATTCTTT 2820
C11_020424      AAATAATAAAGATTCGGATGCATCATAGATATGAAATTTTAAAGGCGCACGCAATTCTTT 118
                ************************************************************

HHV7_RK_U38     TTCGGCTACTTGAGAAAGCGATGGCCATGTCATATCATTTTTGAGAATGTATTGCTTATC 2880
C11_020424      TTCGGCTACTTGAGAAAGCGATGGCCATGTCATATCATTTTTGAGAATGTATTGCTTATC 178
                ************************************************************

HHV7_RK_U38     TTGATAGAACATTCTCGGTTCCGAGTCACAAAGAGTCTTCATTAAGCCGGGTGCTCCATC 2940
C11_020424      TTGATAGAACATTCTCGGTTCCGAGTCACAAAGAGTCTTCATTAAGCCGGGTGCTCCATC 238
                ************************************************************

HHV7_RK_U38     GTGCATTATACCGCGCGGAAAAATCCGGAGAAAAGTAGATTTTGTTTTTTTCTTAGTTCG 3000
C11_020424      GTGCATTATACCGCGCGGAAAAATCCGGAGAAAAGTAGATTTTGTTTTTTTCTTAGTTCG 298
                ************************************************************

HHV7_RK_U38     TACATTCTCTAGATACGGATTAAAGAAGGACACCAGATCCAT- 3042
C11_020424      TACATTCTCTAGATACGGATTAAAGAANGACACCAGATCCATA 341
                *************************** ***************

HHV7U39

HHV7_RK_U39     GTTTACCATGTGCACTTCTTCAATTGGCATTGGCAAAGTTGTAACAGTTCGATCTAAGAA 2160
C12_020314      ---------AAAAAAAAANTNAAN-GANAAAAACAAAGAAAAAAAAAAAAGAAGNANTANA 50
                         *     * **  *  *    *****    ** *     *       *

HHV7_RK_U39     ATACACAGTGCCAACATCGCGATATGTTGTTTGAAATGTAAGTTCTTTTTTAAAAGTTCT 2220
C12_020314      NGGNTNGGTGGNANAATNNCTCTGGATTGTGTTGAATGTAG---TCNNTTNNNGTGNNGG 107
                   ***  *  **  *  *   **** *  ******         **      *

HHV7_RK_U39     GACAGAAAACGTGTGTGCTTCAATCTTTGTTTTGTAAATTATCAAAATACCTTCGGTAGT 2280
C12_020314      GACGANCGCTCNCTGTGCTTN--TCTTTGTTTTGTNNTTTATNNNN----CTTCGGTAGT 161
                ***           *******   *************   ****          **********

HHV7_RK_U39     TTTAATGTTAGAGCCGTAAGACGCGCAAGAAACCTCCCTGTCAAAACGCACTAAATCTGT 2340
C12_020314      TTTAATGTTNGAGCCGTNNGACGCGCNNGAC-CCTCCCTGTCA---CTCGCCTA--CTTG 215
                ********* *******  *******  **  ***********   * * *  *  **

HHV7_RK_U39     CCCGGTGGCAATTGAACAAATTCGAAATGGTAAATGCTGATTGTGTCCAGTCATGACAAA 2400
C12_020314      TCCGGTGGCC-TTGAAC---CCTNNNCTGGTNT-TGCTG-TTTTGTCNGTCTGACCCCTC 269
                 ********  ******           ****   ***** ** ****         *

HHV7_RK_U39     GTCAGCTTCTGTTTGTAGAGATAGCTGTAAGCTAAAAGTTATAAAAACACTCAGGAATAG 2460
C12_020314      NCT---------------------------------------------------------- 272

HHV7U40

HHV7_RK_U40     AAGAGCCAAGTCCAATTTTAAGTGTTGCTTGCATTCATTGAGTTGGAATATATTTTCAAA 1860
D1_020228       ---------------------------------------GAGNNGGAATATATTTTCAAA 21
                                                        ***  ****************

HHV7_RK_U40     ATATTTTTG-TTTGTCATTATCTAATAGAAATTGATGAATTGAAGCATCGAGTAATAAGA 1919
D1_020228       ATATTTTTNGTTTGTCATTATCTAATAGAAATTGGTGAATTGAAGCATCGAGTAATAAGA 81
                ********  ************************ *************************

HHV7_RK_U40     CTTGATTATACATAGCTTTTAATAAAATCTGTAAATAAATGGTTACAGGAGA-TGCACAA 1978
D1_020228       CTTGATTATACATAGCTTTTAATAAAATCTGTAAATAAATGGTTACAGGAGAGNGCACAA 141
                ************************************************  ******

HHV7_RK_U40     TTAAAATTTTGATGACAGAGTTCAGCAAATAATGTCTTAAACAAATGAATTATTACAACA 2038
D1_020228       TTAAAATTTTGATGACAGAGTTCAGCAAATAATGTCTTAAACAAATGAATTATTACAACA 201
                ************************************************************

HHV7_RK_U40     CCGTTTTTCCTGAAATTTGTCATATCTGGAATAAGTGTGACAGGGTCACAAAATCGTAAG 2098
D1_020228       CCGTTTTTCCTGAAATTTGTCATATCTGGAATAAGTGTGACAGGGTCACAAAATCGTAAG 261
                ************************************************************

HHV7_RK_U40     CACTCTAATTCTAAAGCACATTCACTTAACCTTGAACACAAGACACACAATGATTGCAAA 2158
D1_020228       CACTCTAATTCTAA-GCACATTCACTTA--CCTNNACACAACACAACATATG-------- 310
                ************** *************  * *  ****** ***    ***
```

HHV7U41

```
HHV7_RK_U41     TCCTATATCCACAGGCTTTCGATTTCCGTTAACGACAAATTCTTGTAGATTGAATTTTTC 3060
D2_020424       ----------TCAGGCTTTCGATTTCCGTTAACGACAAATTCTTGTAGATTGAATTTTTC 50
                          **********************************************************

HHV7_RK_U41     TCTTGTTTGTTTGCACAATTTATCAAGATTTCCATGGTCTTCTGCTTCTTTTAAAAAAAC 3120
D2_020424       TCTTGTTTGTTTGCACAATTTATCAAGATTTCCATGGTCTTCTGCTTCTTTTAAAAAAAC 110
                ************************************************************

HHV7_RK_U41     ATCTGTACCGTGGAAAAAGAAGCATACAGGCATAAACGATGTGATTTTGGTAATGACTGT 3180
D2_020424       ATCTGTACCGTGGAAAAAGAAGCATACAGGCATAAACGATGTGATTTTGGTAATGACTGT 170
                ************************************************************

HHV7_RK_U41     TCCGCCGTAATTAATAATGGGTGTTTTAACGGTTGGGAAGAAGTCATTTTCAACCGTTAG 3240
D2_020424       TCCGCCGTAATTAATAATGGGTGTTTTAACGGTTGGGAAGAAGTCATTTTCAACCGTTAG 230
                ************************************************************

HHV7_RK_U41     ATTAAGTAACAAGGGTGAGATCACTACTGATTTTCTTTTTTCCATTAACGATAATGTAGT 3300
D2_020424       ATTAAGTAACAAGGGTGAGATCACTACTGATTTTCTTTTTTCCATTAACGATAATGTAGT 290
                ************************************************************

HHV7_RK_U41     CAGAATTTCTATTAATTTTTGTTCTTTCGGCAGAATGTACAGCCATGCTGCGGTACATAT 3360
D2_020424       CAGAATTTCTATTAATTTTTGTTCTTTCGGCAGAATGTACAGCCATGCTGCGGTACATAT 350
                ************************************************************

HHV7_RK_U41     TGGGGCCGAAACAACAGTTTCATTATCATCAGCCAT 3396
D2_020424       TGA---------------------------------- 353
                **
```

HHV7U42

```
HHV7_RK_U42     TGGATCTCGTGCTTCACCGGAATACTTGCGTAAACCAAGGGCTTCCGCGGTGTAGCTAGA 1260
D3_020413       -----------------------------------TAAGGGCTTCCGCGGTGTAGCTAGA 25
                                                   *************************

HHV7_RK_U42     GCCAAAAGGTGCTGTTGGTAAATTGGCGAAAAAAGTCAATGCATGCGTCTGTGCACGCTT 1320
D3_020413       GCCAAAAGGTGCTGTTGGTAAATTGGCGAAAAAAGTCAATGCATGCGTCTGTGCACGCTT 85
                ************************************************************

HHV7_RK_U42     TTCTTCCTGCTTTTCTTCAATGATGAGTTTCAGCTGTTCTAGTTCTCTGAGACTCAAAGC 1380
D3_020413       TTCTTCCTGCTTTTCTTCAATGATGAGTTTCAGCTGTTCTAGTTCTCTGAGACTCAAAGC 145
                ************************************************************

HHV7_RK_U42     GTCTAATTTTGATTCATCAAGCTGCTTAATAGGAATGATGCGTCTGTGAAATTGTTTCGT 1440
D3_020413       GTCTAATTTTGATTCATCAAGCTGCTTAATAGGAATGATGCGTCTGTGAAATTGTTTCGT 205
                ************************************************************

HHV7_RK_U42     AAAGCGGCTGACATACTTGTTAGCAGGTTTATGAACAAGAGTTCGTTTAATGGTTTTCAA 1500
D3_020413       AAAGCGGCTGACATACTTGTTAGCAGGTTTATGAACAAGAGTTCGTTTAATGGTTTTCAA 265
                ************************************************************

HHV7_RK_U42     ACCATATCGTTGCCGCCCTAATACATTCTTTTTAACCCCGCGTGGATACAT 1551
D3_020413       ACCATATCGTTGCCGCCCTAATACA------------------------- 290
                *************************
```

HHV7U43

```
HHV7_RK_U43     CCAAAACGGTGACGTGGGTAAGAGATACTTATTTGTTAGATGTTTAATGATTATAGTCTT 2220
D4_020313       ------------------------------------------------------------

HHV7_RK_U43     GCAAAAAAAGAC--TAAGCTTTTAAATTCTATATTTTCACTTCCTTCTGGAACATGAGAA 2278
D4_020313       GCAGGNCCCGGCGCCAGGCGGCNGNGGGANNCGATTNCACTTCCTTC-GGAACAGGAGAA 59
                ***      * *   * **              ** ********** ****** *****

HHV7_RK_U43     AGTAAAAAAATCTTTATACAGATTTTCTGTTTCTTCCGTATTTAAACAGTGAATAATATCC 2338
D4_020313       AGNNAAAAAT-TCCNNNCAGANNCNNTNNNN---NAAANANTAAACAGGAAANA---TCC 112
                **  ****** *     ****     *              *******  ** *   ***

HHV7_RK_U43     TCAGACTTAATGGGTAATGCATTTTTTACATAATAGCTGAAATTCAAAGATTCTGCGTCG 2398
D4_020313       TCAGACT-AAGGGNNAGCANA-------CANGANAGCGGAAATTCAAAGATTCGGNGTNG 164
                ******* ** **  *    *       **  * *** *************** * ** *

HHV7_RK_U43     CAAACAAATACCGTCTCAATTCTTTTGGAAAACTTGCATTTCTGTGTCTGCATACAAAAA 2458
D4_020313       CAAACAAATACCGTCTCAATTCN---GGAAAACT-GCATTCGNGT---CGCATACAAAAA 217
                **********************    ******** *****   **    ************

HHV7_RK_U43     TATACATTTTTTGATGGTTTATATTTAACGATTATTGGAAAAATTAAATGTTCACTAGGG 2518
D4_020313       TANG-ACANNNNGAGGNCTTA--ATTNACGATTAT--GGAANATTAA--GNTCACNNGGN 270
                **   *      ** *  ***   ** ********  * ** *****  * ****  **

HHV7_RK_U43     TTCTTACATAAGATATTTACTATCACGTGCGCGGGGTCATATTCCGTTGCAAACACAGTA 2578
D4_020313       TC--TACATGA-ATACATNAGGTGGGGT-CAANANGCNACNNTCAGGNNGTCGCGCCGTG 326
                *   ***** * ***  *    *   ** *     *  *   ** *       * * **

HHV7_RK_U43     ATAGTCAT--- 2586
D4_020313       CAGANAGTGCT 337
                  *
```

HHV7U44

```
HHV7_RK_U44     ATGGGAAATTGCTTTGTAAAAAAAATCTCTTCTGACATCTTCAACTCTAACTACCAGATG 60
D5_020413       ------------------------------------------------TCTACCAGATG 11
                                                                 **********

HHV7_RK_U44     CTCTATTCTGAGCTCTCAGAGCAAGAGGATTTACTGGATTTTTTAGAGACAAAATATACA 120
D5_020413       CTCTATTCTGAGCTCTCAGAGCAAGAGGATTTACTGGATTTTTTAGAGACAAAATATACA 71
                ************************************************************
```

```
HHV7_RK_U44     GATTTTGGAATTTTAAAAACCGATATTCTCAACTATGAAAGAGACTCCGAAACATTCAAA 180
D5_020413       GATTTTGGAATTTTAAAAACCGATATTCTCAACTATGAAAGAGACTCCGAAACATTCAAA 131
                ************************************************************

HHV7_RK_U44     ACTTTGTTGCAAGTGTTACCTATCTATAAAAAAACAAAGCTGAGGTATAATTTGATTGAA 240
D5_020413       ACTTTGTTGCAAGTGTTACCTATCTATAAAAAAACAAAGCTGAGGTATAATTTGATTGAA 191
                ************************************************************

HHV7_RK_U44     CGCTGTTTAAATAATTGTCCCCCTCACGTTAAAGATGCATTGATTATTGAAATCATGAAA 300
D5_020413       CGCTGTTTAAATAATTGTCCCCCTCACGTTAAAGATGCATTGATTATTGAAATCATGAAA 251
                ************************************************************

HHV7_RK_U44     GCTAAGAAAATTTTAGAGACTCTGGATGTTGTGTTCATGAAAATTATGATTGGAGAATTT 360
D5_020413       GCTAAGAAAATTTTAGAGACTCTGGATGTTGTGTTCATGAAAA---------------- 294
                ******************************************

HHV7U45

HHV7_RK_U45     TTTTTCACAGGATTTGAATAAACCGTTTTGAGGAATT-CCAAGTAAAGAAAATGGTCGTT 899
D6_020502       ----------------------TCCGTTTTGAGGAATNTCCAAGTAAAGAAAATGGTCGTT 39
                                      ***************  ***********************

HHV7_RK_U45     GATAAACCGTCAATTCAGTGTTCACATATTCATTTAATTGCTTTGCATCGATTTTAGGCC 959
D6_020502       GATAAACCGTCAATTCAGTGTTCACATATTCATTTAATTGCTTTGCATCGATTTTAGGCC 99
                ************************************************************

HHV7_RK_U45     ATGTAAGAGAGCTCTCTTTTGTTTCGCTAAAGTTCCAGATATCATCCACATCCGTTACAG 1019
D6_020502       ATGTAAGAGAGCTCTCTTTTGTTTCGCTAAAGTTCCAGATATCATCCACATCCGTTACAG 159
                ************************************************************

HHV7_RK_U45     TAAACAGGAAAAAGTGCAAGTATCTGGACATAGGCAATTTGTTAAGTGTAATTGTTTCAG 1079
D6_020502       TAAACAGGAAAAAGTGCAAGTATCTGGACATAGGCAATTTGTTAAGTGTAATTGTTTCAG 219
                ************************************************************

HHV7_RK_U45     AGATTTTTTCAGAAATTATTTTCTTGTTTGTTTCAGATCTTCT--GGAGGGTAGACTTTG 1137
D6_020502       AGATTTTTTCAGAAATTATTTTCTTGTTTGTTTCAGATCTTCCTGGAAGGGTAGACTTTG 279
                ******************************************   * **************

HHV7_RK_U45     CAT 1140
D6_020502       CNA 282
                *

HHV7U46

HHV7_RK_U46     ATGACTCTATACAAGATTGTTTCCAAGCCAATTATATTGCTGGCATTTTTTTTTACCCGA 60
D7_020419       --------------------------GCCAATTATATTGCTGGCATTTTTTTTTACCCGA 34
                                          ***********************************

HHV7_RK_U46     GTTGTTTTCACAAATGAAGTCGACGGGGAGGAGCTGTTTTATAAGCCTACTTGTCATTCA 120
D7_020419       GTTGTTTTCACAAATGAAGTCGACGGGGAGGAGCTGTTTTATAAGCCTACTTGTCATTCA 94
                ************************************************************

HHV7_RK_U46     GATACATATGAAATTATTTTAAAAAAAATTTTCTTCTATTTGGATATTGGTCAACACATTT 180
D7_020419       GATACATATGAAATTATTTTAAAAAAAATTTTCTTCTATTTGGATATTGGTCAACACATTT 154
                ************************************************************

HHV7_RK_U46     ATTTTATTATGTTCATTTTCATTATTTTTAAAGTATTGGTGTTTTAAGACTCTC-GCCAA 239
D7_020419       ATTTTATTATGTTCATTTTCATTATTTTTAAAGTATTGGTGTTTTAAGACTCTCAGCCAA 214
                ************************************************** *****

HHV7_RK_U46     AGAAACTGTAAAAGGGTATTAA 261
D7_020419       AGAAACTGA------------- 223
                ********

HHV7U47

HHV7_RK_U47     GGGAAATTCAAAATGTAATGTCCATCGCGGAAAATTATCTAGATAGAGCATATTTATTTT 660
D8_020502       ---TACNNTGGNNANTAACNGTACTCNNGTNAAAT-ATCN-GAT-GAGC---TNTTATTT 51
                   *          ***      **  *  **** ***  *** ****   * *  ***

HHV7_RK_U47     AGAAAGGCAAATATTCTTTTTCTGGGATAGACGGAATGACCAAAATAAAATCTGTTTAAA 720
D8_020502       AGAAAGGCAA-TATTCTTTTTCTGGGATAGACGGAATGACCAAAATAAAATCTGTTTAAA 110
                ********** ************************************************

HHV7_RK_U47     ACAGTTCTTAGAGATGGGTATGTCGCAACTCTGAATATATGTTTTTTCAAAAGACTCATA 780
D8_020502       ACATTTCTCAGAGATGGGTATGTCGCAACTCTGAATATATGTTTTTTCAAAAGACTCATA 170
                *** **** ************************************************

HHV7_RK_U47     ATGAAATTTTGCGAATTTCGGCCACATGGAAAAATTCGTTAAATTTACAAACGAGATTTT 840
D8_020502       ATGAAATTTTGCGAATTTCGGCCACATGGAAAAATTCGTTAAATTTACAAACGAGATTTT 230
                ************************************************************

HHV7_RK_U47     TTCAGAGCAAACACTTAATACAGGAGTCCATATGATCAGATAACTATAAAAAAAGAAAAGA 900
D8_020502       TTCAGAGCAAACACTTAATACAGGAGTCCATATGATCTTATAACTATAAAAAAAGAAAAGA 290
                ************************************  ***********************

HHV7_RK_U47     AATCAACGTGAACACTAGCATTGAATACATTTTGTTTTTCAT 942
D8_020502       AATCAACGTGAACACTAGCATTGA------------------ 314
                ************************

HHV7U48

HHV7_RK_U48     CTTTTCATAAAATAAAGGTAATAAAGACACTGTAATCGTTTTTGTATGCGAAACAGGTTT 1740
D9_020413       ----------------------------------------TTTGTATGCGAAACAGGTTT 20
                                                        ********************

HHV7_RK_U48     GATATTTGTCTGAAATTTTTGTGCA-TAGATTTTCACAGCTTTCTCTGTTGATGGGTTAA 1799
D9_020413       GATATTTGTCTGAAATTTTTGTGCACTAGATTTTCACAGCTTTCTCTGTTGATGGGTTAA 80
                ************************** ***********************************
```

```
HHV7_RK_U48     AAAAACGTGTAATCCGATGTTTATATTGATCAAATGATTCATCGAAATTCACAGAATCAA 1859
D9_020413       AAAAACGTGTAATCCGATGTTTATATTGATCAAATGATTCATCGAAATTCACAGAATCAA 140
                ************************************************************

HHV7_RK_U48     ATAAATTCGAGACAAAAGTATATCCGAATAAGCATTTGGGAATCTGATAAACCCGAGTTT 1919
D9_020413       ATAAATTCGAGACAAAAGTATATCCGAATAAGCATTTGGGAATCTGATAAACCCGAGTTT 200
                ************************************************************

HHV7_RK_U48     CGTTATAATCATGGAAATTAAACGTTATTAAACCAGGTTGAATAATAGT-TTGATTTGTT 1978
D9_020413       CGTTATAATCATGGAAATTAAACGTTATTAAACCAGGTTGAATAATAGNCNTGATTTGTT 260
                ************************************************  *********

HHV7_RK_U48     TTT--TCATTTACGCAAATTGAAGAATTGAGTATATTCCAGTGTTTCCATCCATTTATTG 2036
D9_020413       TTTAANCATTTACGCAAATTGAAGAATNGAGTATATCCCAGTGTTTCCATCCATTTA--- 317
                ***   ********************** ******** *********************


HHV7U49

HHV7_RK_U49     ATGTCTTTAGAATACTTGCCCCCTGTTAGAAGACGTATTGGTCAATATAATCATTTGAGA 60
D10_020413      -------------------------TAGAAGACGTATTGGTCAATATAATCTTTTGAGA 34
                                         *************************** *******

HHV7_RK_U49     ATCTACAAAAAAAATTCTGTTGTTAAAATCCAATTTTGAAAAATTAAATTTTTTTTTGGGC 120
D10_020413      ATCTACAAAAAAATTCTGTTGTTAAAATCCAATTTTGAAAAATTAAATTTTTTTTTGGGC 94
                ************************************************************

HHV7_RK_U49     AATCTTTTTCCAGAAGAACTGCATGATTCAAAAATACATGTATATTTTGAAGTTAGATTA 180
D10_020413      AATCTTTTTCCAGAAGAACTGCATGATTCCAAAATACATGTATATTTTGAAGTTAGATTA 154
                ****************************** ******************************

HHV7_RK_U49     GGGTGCCGTATTCCGGATTGCATTATTGTTTTTAGGCATTTTGGTGAGAAATTATTAAAA 240
D10_020413      GGGTGCCGTATTCCGGATTGCATTATTGTTTTTAGGCATTTTGGTGAGAAATTATTAAAA 214
                ************************************************************

HHV7_RK_U49     ACTTTCCATTGTTATTTTTTTGAATTCAAAACTACTTTTGCCAAATCAAATCTGTTTTCA 300
D10_020413      ACTTTCCATTGTTATTTTTTTGAATTCAAAACTACTTTTGCCAAATCAAATCTGTTTTCA 274
                ************************************************************

HHV7_RK_U49     ATT-CAAAAAAATAGGACACAAAAAATTCAATATTTACAGGG-ACTAAGGCAACTACGAC 358
D10_020413      NTTTCAAAAAAATAGGACACAAAAAATTCAATATTTACAGGGGACTAAGGCAACTACGAC 334
                 ** ***************************************** ****************

HHV7_RK_U49     AAGCAACAGATTATCTACAGCAATTTGTCATCAAAAATGAAAGCTTATGTAAGGTAAACC 418
D10_020413      AAGCAACCAGAAATCACT---------------------------------------- 352
                *******     ***


HHV7U50

HHV7_RK_U50     ACTGAACTTGATGTTTTGCAAGCACATGTACAGACAGAATGCCAGAAATTAAATACAAAT 240
D11_020424      ------------------------TATGTACAGACAGAATGCCAGAAATTAAATACAAAT 36
                                       ************************************

HHV7_RK_U50     TTAAGAGACATTGAAAATGCATTGCTTTTAGAAAATCAAAAAATAATTCCTTCTAGTGAA 300
D11_020424      TTAAGAGACATTGAAAATGCATTGCTTTTAGAAAATCAAAAAATAATTCCTTCTAGTGAA 96
                ************************************************************

HHV7_RK_U50     ACACGTTCTGTCTTGGAAGAGTCTCTCCAGGCTAAAACTGTCACGCAAGTCACAATTACA 360
D11_020424      ACACATTCTGTCTTGGAAGAGTCTCTCCAGGCTAAAACTGTCACGCAAGTCACAATTACA 156
                **** ********************************************************

HHV7_RK_U50     CAAATTGATCCTGCTATTCATTTCACGGAAAATTTCCGACCTGAGATGATAAAAACTTTT 420
D11_020424      CAAATTGATCCTGCTATTCATTTCACGGAAAATTTCCGACCTGAGATGATAAAAACTTTT 216
                ************************************************************

HHV7_RK_U50     TATAATAATACTCAGATGTGGAGCTATACGTTTGGAGCTTGGTTTTACAAATTAAAGCGA 480
D11_020424      TATAATAATACTCAGATGTGGAGCTATACGTTTGGAGCTTGGTTTTACAAATTAAAGCGA 276
                ************************************************************

HHV7_RK_U50     GCTTTCTTTACCGATTCTAAATTAAAAAGAATGCTAAAGTTAACATATGTTGATTCTCTC 540
D11_020424      GCTTTCTTTACCGATTCTAAATTAAAAAGAATGCTAAAGTTAACATATGTTGATTCTCTC 336
                ************************************************************

HHV7_RK_U50     TCTATTACACAGGAATTGCTGTCTATTTCAATTAATGCATTGGAACAAATTACTATTTAT 600
D11_020424      TCTATTACACAGGAATTGCTGTCTATTTCAATTAATGCATTGGAACAAATTACTATTTAT 396
                ************************************************************

HHV7_RK_U50     CCTATGCATGATAATTTAGTGTCCGATTTAGAAGCTGGTTTGTGTCTACTGACTGCTTTT 660
D11_020424      CCTATGCATGATAATTTAGTGTCCGATTTAGAAGCTGGA---------------------- 435
                ****************************************


HHV7U51

HHV7_RK_U51     TTTTTCTCGTTTTTCTTTCTTTGTCTGCTGGTAATAATAGTGGTGAAATTTAATAACAGC 120
D12_020314      ----TCTCGGNTTCNTNNNTTCN--NGCTGG-AATAATAGTGGTAAAN--NAATAACAGC 51
                    *****  **  *   **      ***** ************ **    *********

HHV7_RK_U51     ACCGTTGGTCGAGAATACACATTTAGTACT-TTTTCTGGAATGTTGGTT-TATATTTTAT 178
D12_020314      ACCGTGGTCGGAGAATACACANN-AGNACTGTTTTCTGGAATGNGGGTTANANNNATNNN 110
                ***** *   ************   ** *** *************  ****  *    *

HHV7_RK_U51     TGT-TACCGGTTAAAATGGGAATGTTAACCA----AAATGTGGGAT--GTATCCACCGAT 231
D12_020314      NGAATACCGGTTAAATGGGGAATGTTAACGGCGGGAAATGTGGGANCTGNATCCACCGAT 170
                 *  ************  *************      ***********   * **********

HHV7_RK_U51     TACTG-TATAATCTTAATGTTTCTGAGTGATTTTTCATTTATCTTTTCGTCCTGGGCGTT 290
D12_020314      TACNNGTATNATCTTNATGTTTCTGAGTGATTNTC--ATTATCTTTTCGTCCTGGGCGTT 228
                ***   *** ***** ***************** *    **********************

HHV7_RK_U51     AACGCTGTTGGCACTGGAACGTATCAACAATTTTTCTTTTTCTGAGATTAAGGTAAACGA 350
```

```
D12_020314      A-CGCTGTTGGCT---------------------------------------------- 240
                * **********

HHV7U52

HHV7_RK_U52     TAATTTTTTCGAGAGCTGAGCTGCATTTTCTTCACTGAAGATATGTTTAAGGATTTCT-- 478
E1_020228       -------------TGCTGAGCTGCATTTTCTTCACTGAAGATATGTTTAAGGATTATAAA 47
                            ****************************************

HHV7_RK_U52     GATCCAGATCTTTGTTTAATCATGCAAAAGATCTTAAACCAAAATTCATCAGGTACCTTT 538
E1_020228       GATCCAGATCTTTGTTTAATCATGCAAAAGATCTTAAACCAAAATTCATCAGGTACCTTT 107
                ************************************************************

HHV7_RK_U52     CGATTCAGAA-CCACTAAATCACTCACTCCGATATTGTGCAAAGTCTCTCTCAACAGTAA 597
E1_020228       CGATTCAGANACCACTAAATCACTCACTCCGATATTGTGCAAAGTCTCTCTCAACAGTAA 167
                *********  **************************************************

HHV7_RK_U52     GATTTGGATGCCGTGAGAAAACATTAGGGTAC-ACACATTCTGGATGATCAAAATCTCTT 656
E1_020228       GATTTGGATGCCGTGAGAAAACATTAGGGTANCACACATTCTGGATGATCAAAATCTCTT 227
                *********************************  ***************************

HHV7_RK_U52     CTTGTTTTAGATTAAGTAAAGATTCACCGCTTACCAATTTCTTAGTGATATGTAATATTA 716
E1_020228       CTTGTTTTAGATTAAATNAAGATTCACCGCTTACCA-TTTCTTAGGTANATGTTAATTTA 286
                *************** * ******************** ********  * **** *  ***

HHV7_RK_U52     AATTGGAGGTTGAGTTGTTCGAGATTATGTATTGACCTACTTGTGTCAT 765
E1_020228       NGGNGGAAGT-GCGTTGTTC-AGATCGAGNNT-GGGTTGCCGTCGC--- 329
                   *** ** * ******* ****   *  * *   * *    *

HHV7U53

HHV7_RK_U53     ATGGAAACTGTGCTAGTGGCAGGCTTCCTATGTGTGTATGATGATAATGATATAAATGAT 60
E2_020424       --------------TGTGGCAGGCTTCCTATGTGTGTATGATGATAATGATATAAATGAT 46
                              **********************************************

HHV7_RK_U53     AACTTTTATCTGCCAAGGAGGACGATACAAGAAGAAATCAATTCCGGAAATGGTTTGAAT 120
E2_020424       AACTTTTATCTGCCAAGGAGGACGATACAAGAAGAAATCAATTCCGGAAATGGTTTGAAT 106
                ************************************************************

HHV7_RK_U53     ATTCCATTGAATATAAACCACAATGAAAATGCCGTTATAGGCACAGTCTCTTCTTTAAGT 180
E2_020424       ATTCCATTGAATATAAACCACAATGAAAATGCCGTTATAGGCACAGTCTCTTCTTTAAGT 166
                ************************************************************

HHV7_RK_U53     GATTTACAGCACGGTTTGTTCACGGTTGCCCGTGTTCAATCAAAGGAATTCCTTACAATA 240
E2_020424       GATTTACAGCACGGTTTGTTCACGGTTGCCCGTGTTCAATCAAAGGAATTCCTTACAATA 226
                ************************************************************

HHV7_RK_U53     ATTAAGAAAATAGCTGTAAAATCTAAGCTGATAACCAACACGGAAGAAAAAACTCTGCCA 300
E2_020424       ATTAAGAAAATAGCTGCAAAATCTAAGCTGATAACCAACACGGAAGAAAAAACTCTGCCA 286
                ********************************************************

HHV7_RK_U53     CCAGATCCCGAAATAGAGTGTTTAAATTCAATTTTCCCAGGTTTGTCATTATCGAACAGG 360
E2_020424       CCAGATCCA---------------------------------------------------- 295
                ********

HHV7U53.5

HHV7_RK_U53.5   ----------------------------------------------------ATGAATGGCTC 11
E3_020413       CCCCGGGNCCCCATTGGGGGCGGGCCCCGGGCCGGNGGNANTTTCGGGAATTTTAGGCTT 60
                                                                   **    ****

HHV7_RK_U53.5   AGCTACTGGCAACTCAGCTAGTCTTCCCGCTCCTCAGTTTTCC------GAATGTG---- 61
E3_020413       AAN-GTTNTTTTTNCNNGCGGNCTTCCCCTTCCNAGGTTTTTTTTCCCGGAAAATGGTGG 119
                *     *       *     * ******  ***   *****         ***  **

HHV7_RK_U53.5   ----TTTTTTTACC------GAAAGATAC--ATTTT--GCTCCTT--ACTGAA-TGCT-- 102
E3_020413       TTTTTTTNTTTNNCCCGGAANAAGAATANCAATTTTTGGCTCCCTTAACTGAAATGCTTN 179
                    *** ***  *       **  ***   *****  ***** *  ****** ****

HHV7_RK_U53.5   ACAGCTGGGG-CCCAAAATAAAAA--CGTAACGCCAGCTGCTCCGATATTTAAAACTGAT 159
E3_020413       ACAGCTGNGGGCCCAAAATANAANAACGTAACGCCAGCTGCTCCGATATTTAAAACTGAT 239
                ******* ** ********* **   ************************************

HHV7_RK_U53.5   GAATATATAACTCCGTACCCAGAATCTCTGAGCAGAATGGATTATGGAAATAGAATGAAT 219
E3_020413       GAATATATAACTCCGTACCCAGAATCTCTGAGCAGAGTGGATTATGGAAATAGAATGAAT 299
                ************************************ ***********************

HHV7_RK_U53.5   TATCACATACCACCTCCATATTGGTATCCCTCTATGCCTGGATTTAATTATAAATCTTAT 279
E3_020413       TATCACATACCACCTCCATATTGGTATCCCTCTATGCCTGGATTTAATTATAAATCTTAT 359
                ************************************************************

HHV7_RK_U53.5   CGTGGTTCCCAAAAACGGTGCGCTCCAACAGATTCGGATGATGAAATGAGCTTTCCAGGA 339
E3_020413       CGTGGTTCCCAAAAACGGTGCGCTCCAACAGAATTCGGCAGTAGTA-AANTCACNTAGNG 418
                ********************************** *  *    *   *  *       **

HHV7_RK_U53.5   GATCCTGACTATACGACAAAGAAAAAGAAGCGGTATAGAGAAGATGACGATCGTGAACTC 399
E3_020413       TAANTTGCNNGGCNNNCCCGATCGGACAAGNTNNGCCTTNAGGATNCN--TNANTACTTC 476
                 *   **         *       * ***          * ***     *    *  **

HHV7_RK_U53.5   ACTAAAGACAAAAATGATATTAAAGAACTAGTTGATGCAATAGGGATGCTGCGTCATGAA 459
E3_020413       CNTGTGAGCGGGNTCNNGG--GNGGGACC------------------------------ 503
                  *     *               * **

HHV7U54

HHV7_RK_U54     CTCCATGCGCTCAACTATTTGGACAGGAAG-TGATACCAAATTAATGCTAGGAAGAAAAA 1079
E4_020313       -----NGCGCTCAACNANN-GGACAGGAAGGNGATACCAAATNAATGCTAGGAAGAAAAA 54
                     ********* *   **********  ********** ******************
```

```
HHV7_RK_U54     AGGCAAAAAGATTAATAATTAAAG--GAAATTGTTTTAAATCTAAATCTACATCAGATAT 1137
E4_020313       AGGCAAAAAGATTAATAATTAAANAANNAATTGTTTAAATCTTAAATCTACATCAGATAT 114
                ************************      ******** **   ******************

HHV7_RK_U54     ATTTTGAACAGCCAAGGTTACATTGCCACAATCTCTTAGATCGATATACGTAAAGTAGGC 1197
E4_020313       ATTT-GAACAGCCAAGGNTACATTGCCACAATCTCTTAGATCGATATACGTAAAGTAGGC 173
                **** ************ ********************************************

HHV7_RK_U54     TGCTGTAA--AGCTTCCTTCATCTCTGCACATACACACAAATGTTGATTGTTGCGA---A 1252
E4_020313       TGCTGNTNGGAGCTTCCTTCATCTCTGCACATACACACAAATGTTGATTGTTGCGGGGGN 233
                *****      ************************************************

HHV7_RK_U54     AGTATGCTTAGCCC-TGTCCTTAAGCATCTTATTTCACGAGGTTTTAAAAGGATTGGTGC 1311
E4_020313       AGTATGCTTAGCCCCTGTCCTTAAGCATCTTATTTCACGAGGTTTTAAAAGGATTGGTGC 293
                ************** ************************************************

HHV7_RK_U54     TGTCAATTTTAGTTGTAGAAAATATGAAGACCAGACCACATAACAGGATTCCAT 1365
E4_020313       TNNCAATTTTAGTGNTAGAAAATA-GTAAACCNNACCANANNACGGTATCTCTA 346
                *  **********  ********* * * ***  **** *  ** * **  *
```

HHV7U55A

```
HHV7_RK_U55A    TGTTCCTAAACCCTTTTCAAG-ATAGGCAAATTG-ATGAAATCTAAGAGGATCCATATTT 958
E5_020413       --------TACCCTTTTCAAGGATAGGCAAATTGGATGAAATCTAAGAGGATCCATATTT 52
                        ************ ************ *************************

HHV7_RK_U55A    TCTTTAAAGTAACAATTATCTTGTATCAG-CAGGGATGACGTAATGTTAAAAATTGGCAG 1017
E5_020413       TCTTTAAAGTAACAATTATCTTGTATCAGACAGGGATGACGTAATGTTAAAAATTGGCAG 112
                ****************************** ******************************

HHV7_RK_U55A    GGCAAAAATATTGACAAAGTTTATCGATTGTATCATAGGAAAAACATTAATTTTGATCAC 1077
E5_020413       GGCAAAAATATTGACAAAGTTTATCGATTGTATCATAGGAAAAACATTAATTTTGATCAC 172
                ************************************************************

HHV7_RK_U55A    TTTTTTGTTTTTTAGAACGTGACCCACCAAAGTTACGCATTCTGCATTGTTTTCAGTTCC 1137
E5_020413       TTTTTTGTTTTTTAGAACGTGACCCACCAAAGTTACGCATTCTGCATTGTTTTCAGTTCC 232
                ************************************************************

HHV7_RK_U55A    ATCGGTAATTAGCAACAGTACAGGTCTTTCGTGCGAAAGTATGCAGGTAATCTCTGCGGG 1197
E5_020413       ATCGGTAATTAGCAACAGTACAGGTCTTTCGTGCGAAAGTATGCAGGTAATCTCTGCGGG 292
                ************************************************************

HHV7_RK_U55A    AGAAAAATTGA-GTTCTGGCTGCAAATCATTAAAATCCACAGTAAACTGAATTAATTGCG 1256
E5_020413       AGAAAAATTGAAGTTCTGGCTGCAA---------------------------------- 317
                *********** *************
```

HHV7U55B

```
HHV7_RK_U55B    CGTCACTATGACCTTTTCGTTTTTGAATCTTATCAGACTAGGAAAAACTCCCATCTGCTC 960
E6_020424       ----------------------------------------TGGAAAAACTCCCATCTGCTC 21
                                                        **********************

HHV7_RK_U55B    TTGTTCTTTAACATCATGGTTTGGGATGATGACGATAGAGCTGTTAATTTTTAACAACGG 1020
E6_020424       TTGTTCTTTAACATCATGGTTTGGGATGATGACGATAGAGCTGTTAATTTTTAACAACGG 81
                ************************************************************

HHV7_RK_U55B    AAATCCAAAAAACAAGACTTTTGATGGATAAATGTGTCTATTCAGATTTTTTAAATGTAG 1080
E6_020424       AAATCCAAAAAACAAGACTTTTGATGGATAAATGTGTCTATTCAGATTTTTTAAATGTAG 141
                ************************************************************

HHV7_RK_U55B    TTTTAGTTGTTGTTCTTCCATTAAAGGACGAGCAGAAAAAAAGGAGTCAGAATTCCTTGA 1140
E6_020424       TTTTAGTTGTTGTTCTTCCATTAAAGGACGAGCAGAAAAAAAGGAGTCAGAATTCCTTGA 201
                ************************************************************

HHV7_RK_U55B    ATTTATTTCAGCGCTCAAAAGTAAAACTGGTTCAGTTTTTGATTGGACATACTCGTCAAA 1200
E6_020424       ATTTATTTCAGCGCTCAAAAGTAAAACTGGTTCAGTTTTTGATTGGACATACTCGTCAAA 261
                ************************************************************

HHV7_RK_U55B    TAATAATGTAACATTATCTTTACCCTGGAATTCTTCTTTGTATGGAATTTCAATCTGAAC 1260
E6_020424       TAATAATGTAACATTATCTTTACCCTGGAATTCTTCTTTGA---------------------- 302
                ****************************************
```

HHV7U56

```
HHV7_RK_U56     GTCTGTAGGGATCAGTAATGGCAGAATCAGCTCC-CAGTTAGGTAAGACAAGTAATTGTT 539
E7_020419       ------------------TGGCAGAATCAGCTCCTCAGTTAGGTAAGACAAGTAATTGTT 42
                                  ***************** ****************************

HHV7_RK_U56     GCGAAGGCTTCAGAAATGGTGGCAGTAAACTCAATTGATCTCCCTTTTCCAGAGGAAATG 599
E7_020419       GCGAAGGTTTCAGAAATGGTGGCAGTAAACTCAATTGATCTCCCTTTTCCAGAGGAAATG 102
                ******* ********************************************************

HHV7_RK_U56     GTCCGGTATTTTTGATTTCATACTGGTGTCCAATTTTTGGACGTATTAATATTAATTGGT 659
E7_020419       GTCCGGTATTTTTGATTTCATACTGGTGTCCAATTTTTGGACGTATTAATATTAATTGGT 162
                ************************************************************

HHV7_RK_U56     TTCCTTCTACTTTTTGTAAGATTGTTAACACCATTGTTTTTTAAGATATCCCTTATTTGAA 719
E7_020419       TTCCTTCTACTTTTTGTAAGATTGTTAACACCATTGTTTTTTAAGATATCCCTTATTTGAA 222
                ************************************************************

HHV7_RK_U56     TATAATCGGTGGTTGTAGATAAGACCGTGTAAAGTCCTAAATTTGTGTTGCC--AATTAA 777
E7_020419       TATAATCGGTGGTTGTAGATAAGACCGTGTAAAGTCCTAAATTTGTGTTGCCCCAATTAA 282
                ************************************************  ******

HHV7_RK_U56     ATGACGGTGCGACGG-TATT-GGGATGACGGCATTAGTGAGTTTGCATAAAGTTCCAATG 835
E7_020419       ATGACGGTGCGACNGGTATTTGGGATGACGGCATTAGTGAGTA----------------- 325
                ************* * **** **********************
```

HHV7U57

```
HHV7_RK_U57     TGTGTTTTTAAAGGTGTTTTCTAAATGCGCTAAACAAGCAGCACTTAATTCAAACGATAT 3660
E8_020425       ----------------------GAGGGGGNNCGCGCCAANGACGNNANNGATCAGATAG 37
                                       * * *     *   * **   *   *   ****

HHV7_RK_U57     ATTGATAGGATGCTTTTCTGAATACTTGGTAACCAGTACGGTTGTTTCTTTAGGCGCAGT 3720
E8_020425       GGGAGCNGCNCGNATNGAGGTGCNNCNCGCANNNNGCGGGGGNNNCTATT-GGACANNAC 96
                      *   * *    *         * *    *   **     * **  * *

HHV7_RK_U57     TACGTCGTTTCCTGTAGCGACTCTAGGTAATTGTATATAAAACAGAATTTTTCCCAGTGA 3780
E8_020425       NNTNN----TCCTNNA-CGACTC-AGGTA--TGGATANNNAA-ANNANTTCCNNNACAAG 147
                     ****  * ****** *****  ** ***   ** *  * **      *

HHV7_RK_U57     CATTTTGTCTAGGTCGTTGAAACGGATAACATTTGCAGCAACCGCTATGGACGTATGAAA 3840
E8_020425       CGTNGAGT----GAC----AAACGGATA-------CATTGAGNNCCGTAGACGTANNNAA 192
                * *   **    * *    *********        **   *   *  * ******   **

HHV7_RK_U57     AAAATCTATCCATTCAGTCCGATTACAGTAAATCCCAAGTAATGCTTCGAAACTTATGTT 3900
E8_020425       NATA--TATCCATTCAGTCCGATTACAGTNA-TCCCAAGTAATGNTTNNGNNCNTANGN- 248
                 * *  ************************ * ************ **      * ** *

HHV7_RK_U57     ATAACGA-TCAGAGTCGTCACCGTAATACAATCTTAAGTTTTCAAAAAGTTGTTCAGCAG 3959
E8_020425       ATAACGGGTCAGAGTCGTCACCGTA-TACA---TCTAGTTCATATGTACNTTGTTANTAT 304
                ******  ******************** ****   *  ****   *   *  *  * *  *

HHV7_RK_U57     TTTGTGTTCTGATATCATCAAACACGTTTGGAGAAACATCTAGTTTTGGGAAAATTTCAG 4019
E8_020425       AAAGNAATTNTTTANTT-------------------------------------------- 321
                   *   *    **
```

HHV7U58

```
HHV7_RK_U58     ATGATAGACTCAATATCTGAAGAAACTTTAATAGTAAAGAGCTACACCGTTAATCACTGT 60
E9_020308       ------------------------------------------------------------CCCG 4
                                                                             *

HHV7_RK_U58     GCTAAAAATGTTCCAGTGTTTATTAACTCCTATGATTTAACCGCAGAAGTGGCCAAAAAT 120
E9_020308       GCNGN--ATGTNNTGGTACTA------CCCNATGATN--ACCGCG--ANTGGCCAANT-- 50
                **     ****    **  *        ** *****   *****   * *******

HHV7_RK_U58     GAAGACGTGCGATTAGCGCGACAAGTTCAAATTTCATTAGAG-----AAAATAGATGAAG 175
E9_020308       -GAGACGTGCGATTAGCGCGACGAGTTNCGNNTTTCTTAGACCCCNCNNNATAGATGAAG 109
                  ******************** ****     **  *****          **********

HHV7_RK_U58     TTATAGAATCAATTTTTTCTGCGTCTGGTCCTAGCGTTGAAAATGTAAAAGATCAGGCAA 235
E9_020308       TTATAGAATCCCTTTTTTCTGCGTNTGGTCCTAGCGTTGAN--TGTANAT--CCNGGCN- 164
                **********  ************ ***************   **** *    * ***

HHV7_RK_U58     AGTTTGCTTTGTGTCGTTTACTGCTTGGTCCTGTGA-GTATTCCGTG---------CTAC 285
E9_020308       -GTTTGCTTTGTGTCGTTTACTGCTTGGTCCTGTGAANNATNATNCAACAGATAGCCTAC 223
                 ************************************   **                ****

HHV7_RK_U58     TGCGAAGAATGGGATGTCAATTTTTATCTGACAAAATGTAGTTATAATTGCGAAGGCCCG 345
E9_020308       TGCGANGANTGGGATGT-ATTTTTTCTCTGACC--------------------------- 255
                ***** ** ******** * ***** ******
```

HHV7U59

```
HHV7_RK_U59     ATGACAGCAAAACAAATTTTTCCAGTTTGGAAAGACTGTACACTTTACGTTAACAATGAG 60
E10_020502      --------------------------------TAACCNNGGANGNTATGCAA--NGGGAG 26
                                               **     *   ** *  *     ***

HHV7_RK_U59     ACAGCAACTGTGCATG-AAATCTTGAATTCCGATTTAAGTGAACTGTTACAGTTAAAGAC 119
E10_020502      NNGACTACTNTGCATGGAAATCTGGAATTCCGATTTAAGTGAACTGTTACAGTTAAAGAC 86
                    * *** ****** ****** ************************************

HHV7_RK_U59     GGAATTTGTATCTATGACAGACCTATGTGTTTACATTACTGGATGTATAAATCAGAATAT 179
E10_020502      GGAATTTGTATCTATGACAGACCTATGTGTTTACATTACTGGATGTATAAATCAGAATAT 146
                ************************************************************

HHV7_RK_U59     TTCCAGCATCACGATATATTGGCATGCTTACAGTGAAGTAATTTATGCTTTAACTGGAAT 239
E10_020502      TTCCAGCATCACGATATATTGGCATGCTTACAGTGAAGTAATTTATGCTTTAACTGGAAT 206
                ************************************************************

HHV7_RK_U59     TATACACTGTGAAAAGATTTCTATTGAATGTGGAATTAAATCCACGGACAATAACATTTT 299
E10_020502      TATACACTGTGAAAAGATTTCTATTGAATGTGGAATTAAATCCACGGACAATAACATTTT 266
                ************************************************************

HHV7_RK_U59     GTATGAAAAGCCCAAACTGTTTTTACTTCGAGAAAATTTAGCACCAACTGAATTAAGGTG 359
E10_020502      GTATGAAAAGCCCAAACTGTTTTTA---------------------------------- 291
                *************************
```

HHV7U62

```
HHV7_RK_U62     ATGAACGGAGTTTTAAACGATATAAAGACTGAGTTTTTATGTAATACTAAAACAGATCTT 60
E12_020425      ------------------------------------------------------------

HHV7_RK_U62     TTAACGTTGATACAAAAAATTTGTCTGAACTGTGATTTCATTCTTGAACCGGTAGAATCT 120
E12_020425      ---------------TAAATTTGTCTGAACTGTGATTTCATTCTTGAACCGGTAGAATCT 45
                               ********************************************

HHV7_RK_U62     TTTCCTAAAAAAACCGAGTTGGTTGCGGTGATGTATGATACGCTCGCGGTGGAAATTTTT 180
E12_020425      TTTCCTAAAAAAACCGAGTTGGTTGCGGTGATGTATGATACGCTCGCGGTGGAAATAT-- 103
                ************************************************************ *
```

HHV7U63

```
HHV7_RK_U63     ATGAACAAAAAAAAGATGGACTTGCCTAAATGTCAGTCAA-TCACAG-TTGCTTGTGAAG 58
F1_020228       ----------------------------TATGTCAGTCANATCACAGGTTGCTTGTGAAG 32
                                            **********  ****** ************

HHV7_RK_U63     GAGAATGTTCGCAAATGTACAATTTGCATAATCCTCTGACATTTGAAATGGGTTTGGGAA 118
F1_020228       GAGAATGTTCGCAAATGTACAATTTGCATAATCCTCTGACATTTGAAATGGGTTTGGGAA 92
                ************************************************************

HHV7_RK_U63     ACATCTTTATATGTGTTCGGTGTTTTAAGATACATTTTTGCAATATGCTAGAAGACTGCA 178
F1_020228       ACATCTTTATATGTGTTCGGTGTTTTAAGATACATTTTTGCAATATGCTAGAAGACTGCA 152
                ************************************************************

HHV7_RK_U63     ACCTGATAAATA-CTCATGAAGGATGCGTGTGTTCAAAAACCGGGCTTTTTTATAACGGA 237
F1_020228       ACCTGATAAATAGCTCATGAAGGATGCGTGTGTTCAAAAACCGGGCTTTTTTATAACGGA 212
                ************ ************************************************

HHV7_RK_U63     TGGATGCCAGCCTATTCACATACCTGTATGGAACCTACTGAAGAGCCAAATATGGAGACC 297
F1_020228       TGGATGCCAGCCTATTCACATNNCNGTATNNNNCCTNCN--AGANNCNAN-ATGTNNACC 269
                ********************  * ****    *** *   ***  * *  ***   ***

HHV7_RK_U63     GTTAATGTAGTTGTAGTGCTGTTATCATACGTTTACAGTTTTTTAATACAAAATAAAGCC 357
F1_020228       GCTACTACN--TGGNACTTNGCTGCCCCNTGCCGTG------------------------ 303
                * ** *     **       * *  *    *
```

HHV7U64

```
HHV7_RK_U64     GGGATCTTTTTTCAGTGTGTATTTTTTAAAGAAAAGCTTGTATATACGGAATGGACAAAA 180
F2_020425       ----------------------------------TGCTTGTATATACGGAATGGACAAAA 26
                                                  **************************

HHV7_RK_U64     ATAAAGTTTACTTATGTGTTACATGATCTTGT-AATTTCTCAAATCTTTAAGAATGCCTG 239
F2_020425       ATAAAGTTTACTTATGTGTTACATGATCTTGNTAATTTCTCAAATCTTTAAGAATGCCTG 86
                ********************************  ****************************

HHV7_RK_U64     TATTAAAGAAGTAATACATGGGGCATTAATTCTTTCAGTTCCCATAAATATTGATAACCT 299
F2_020425       TATTAAAGAAGTAATACATGGGGCATTAATTCTTTCAGTTCCCATAAATATTGATAACCT 146
                ************************************************************

HHV7_RK_U64     ACATTTTGATACAGATATTTTAATTCTAAAAATTATTTACCCACATTTTTTGCACGATGA 359
F2_020425       ACATTTTGATACAGATATTTTAATTCTAAAAATTATTTACCCACATTTTTTGCACGATGA 206
                ************************************************************

HHV7_RK_U64     TATTGTCATAAAATTATCGGAAATTTTGTCTGGAGCACCTCGCATACAAAAAACAGTGGA 419
F2_020425       TATTGTCATAAAATTATCGGAAATTTTGTCTGGAGCACCTCGCATACA------------ 254
                ************************************************
```

HHV7U65

```
HHV7_RK_U65     TCTCCTGTTTTCGCTGTTGATAAATTTAAAGATTGCTTGCACTGTAATATAATTATTTTA 180
F3_020313       --TCTCCGTNTCGCNGNNA--TAATTNNAAGATTGCT-GC-CTGTAATATAATTATTTAA 54
                  **    * **** *      ****  ********* ** ***************** *

HHV7_RK_U65     AAAAAGAATTTGGATTTTGTT-TTTAGCTTGGCCATAAACGGAATACATGCCGGACAGTT 239
F3_020313       AAAAGAANTTTGGATTTTGTCCTTTAGCTNGGCCATNAACGGAATAC-TGCCGGACAGTT 113
                ****  * ************  ******* ****** ********** ************

HHV7_RK_U65     TGCAACAAATTCCATTAAACTAAAAAAAATTATAATAACAAACGATTTGGTGTACTACAT 299
F3_020313       TGCAACAAATTCCATTAAACTAGAAAAAATTATAATAACAA-CGATTTGGTGTACTACAT 172
                ********************** ****************** ******************

HHV7_RK_U65     ATTAGAATTGGGATCTTTAACGGTAA-CGGATTTACATTTCATTCCAAAAT--ACAATAG 356
F3_020313       ATTAGAATTGGGATCTTTAACGGTAAACGGATTTACTTTCNTTCCNNNTNNTTNCNNTAG 232
                ************************** ********* **   * *         *  ***

HHV7_RK_U65     TGAACATGTGCTGAATGTGCGACCTATTACGCCAAATCTAATTTATGATACTTGTTCAAT 416
F3_020313       TGAAACTGTGCTGAATGTGCGACCTATTACG---------------------------- 263
                ****  *************************
```

HHV7U66

```
HHV7_RK_U66     TCGATCTGTAAAAACATGGGATTTCTGTAAAAACTTGTTTGTAGCTCTAGATGTTTTTGA 4800
E11_020314      -------ATNGAAAAACAGGGATTCGANNANGNG--------AGNNNTAGANGTTTTTGA 45
                      *  *** *  **  ***    *             **   **** ********

HHV7_RK_U66     TATTG-TACTTCTGTATTTTTCTC-TATGATTGCAGTAAATTTGAGCTTTGACAAAG--C 4856
E11_020314      TATTGGTACTTCTGNGNNTTCNNNANANNNATTGCAGTAAATTTGAGCTTNGACAAATGCT 105
                ***** ********    **       *  ********************* ******

HHV7_RK_U66     GCAGT-TCAAGGGG---TCGCATATTG-CAATTTTAGATTTTACGTGCCGTTGGCGATC- 4910
E11_020314      GCAGTGNCAATGGGGTGCTGCATATTGGCAATTTTAGATTTTACGTGCCGTTGGCGATNC 165
                *****  *** ***     ******** ********************************

HHV7_RK_U66     ACAAAAGAGATATAAAGGTTTAACAT-GCCTGCAATAT-GCATGTGTAAAGCCAAGTTC- 4967
E11_020314      ACAAAAGAGATATAAAGGTTTAACATTGCCTGCAATATTGCATGTGTAAAGCCAAGTTNC 225
                *************************** ************ ********************

HHV7_RK_U66     -TGGAGTTA----ATATAATAAATCGT-TTTTGGCAAAATATCGCGCTGTTTGGAAAAGT 5021
E11_020314      TTGGAGNTAGNAGNTATAATAAANNGTCTTTNNGCAAAATATCGCGCTGTTTGGNAA-GT 284
                 ***** **      **********  ** ***  ********************** ** **

HHV7_RK_U66     TGAAGAAATCTTTACGTCCTGTTCATGTTTCCAAATTATAGACTGATACGCTTTCTGAAT 5081
E11_020314      TGA-GAATNCTTTACN-CCTGT-CATNTTCC--AATTATTTTTTTNTNNTATTTNTNTTT 339
                *** ***  ******  ***** *** ** *  ******    *  *    *** *   *

HHV7_RK_U66     TGCATCTATATCGCATGATCGCAACAT 5108
E11_020314      TNTGCATT------------------- 347
                *    *
```

HHV7U67

```
HHV7_RK_U67     CATGGACACAGATATTGCTCTAGCTGCAATTTATAAAGAAACGACTAAATTAAATGAAAA 60
F4_020313       --------------------GANGNGCAANANAAAGGAAACNGACTAAATTAAAT-AAAA 39
                                    *   ****   * *   **  ************** ****

HHV7_RK_U67     GGATGCTAAAATTTTCT---CGGAGGCAGTGCAGACCGCACTAACTGTGTGTAAAGCAAC 117
F4_020313       GGATGCTAAAANTNCTCGGANNAAAANGNNGCAGACCGCACTAACTGTGNGTAAAGCAAC 99
                *********** *           *       ******************** **********

HHV7_RK_U67     C-GCTCCTAATACACGTCTAAAACTCG-TTGAAACACCAACTAATAACTTCTTACTAGTA 175
F4_020313       CTGCTCCTAATACACGTCTAAAACTCNGTNGAAACACCAACTAATAACTTNN-ACTAGTA 158
                * ************************  * ********************   ******* 

HHV7_RK_U67     ACAAATGTTGTTCCATCAGAAACTTCGAAA-GCAACGACTGAAG-CAAATCTTAATATTG 233
F4_020313       ACAAATGTG--TCCATCAGAAACTTCGAAATGCAACGACTGAANGCAAATCTAATTNNNG 216
                ********   ******************** ************  ******* * *   * 

HHV7_RK_U67     ATGCA------GCGTTGGAAAAACTGGCGTCTTCCTTTAATACAGCGGTACCTGTAAAAT 287
F4_020313       GNNNATTGCATGCGCCNGGAAAACTTN-GNGNTTCCTTAATACAGCGTNACCT--TGAAT 273
                    *       ***   * ******   *   * * **********  ****    ***

HHV7_RK_U67     CATCCAAAAAGTATTTGTTGCAAAATGTGAGAAAAATGACCAGTGAGAACATCGCTCTAA 347
F4_020313       CATCCAAATGATNT-------------------------------------------- 287
                ********   * *
```

HHV7U68

```
HHV7_RK_U68     ATGTCACTGCATGAATTAATAAAACAAACTATGTCCAAAAATTTAGAAAAAAAACATTAT 60
F5_020425       -TGTCACTGCATGAATTAATAAAACAAACTATGTCCAAAAATTTAGAAAAAAAACATTAT 59
                 ************************************************************

HHV7_RK_U68     GAGTTGTTAAAATTAAAACTTGGTGAAGATCATCCTCTTAGCGTTCGACAGCAAATTCAC 120
F5_020425       GAGTTGTTAAAATTAAAACTTGGTGAAGATCATCCTCTTAGCGTTCGACAGCAAATTCAC 119
                ************************************************************

HHV7_RK_U68     GCTCTCAATCAAAATCTTGTATCAGAAAATCTCGAACAGTCCCAGATAATTACTTCTTTG 180
F5_020425       GCTCTCAATCAAAATCTTGTATCAGAAAATCTCGAACAGTCCCAGATAATTACTTCTTTG 179
                ************************************************************

HHV7_RK_U68     ACAAAAATGTTAAAGGATCAAAAGCTGCAGCTGAAAGCGCAAAGGAAAAATGCTGCTCAG 240
F5_020425       ACAAAAATGTTAAAGGATCAAAAGCTGCAGCTGAAAGCGCAAAGGAAAAATGCTGCTCAG 239
                ************************************************************

HHV7_RK_U68     CTAGAATGTGTAGATTTGGATGACATTTTGGATACGGCAGCGGAAGTGAAATCCGTCACC 300
F5_020425       CTAGAATGTGTAGATTTGGATGACATTTTGGATACGGCAGCGGAAGTGAAATCCGTCACC 299
                ************************************************************

HHV7_RK_U68     GACAATATAAAAGAAACTTTACTGGCCGGATTAGAATCAGACTAA 345
F5_020425       GACA------------------------------------------ 303
                ****
```

HHV7U69

```
HHV7_RK_U69     ATGGAGCAGCTTAAGACACCCCAAAATCAAAAAACACGTCCAAGAAATAT--GCTTCCTA 58
F6_020425       ----TGCAGCTTAAGACACCCCAAAATCAAAAAACACGTCCAAGAAATAAAGNCTTCCTA 56
                     ********************************************    *******

HHV7_RK_U69     AAAAAAAAGGAAAAGAACTTAAAAAAAAGGCCTTGTAAAGTAAAACGTAAATTATTTGGTT 118
F6_020425       AAAAAAAAGGAAAAGAACTTAAAAAAAAGGCCTTGTAAAGTAAAACGTAAATTATTTGGTT 116
                ************************************************************

HHV7_RK_U69     CCGAAAACATCAGACCTAACAAAAAAATACCTCTGGCTTCAGACGTGGATAACGAATTGG 178
F6_020425       CCGAAAACATCAGGCCTAACAAAAAAATACCTCTGGCTTCAGACGTGGATAACGAATTGG 176
                ************* ************************************************

HHV7_RK_U69     AAAAAAAACGGGGCTCGATGATACGAAAACGGTCTGAGACGGACTTATGTCCAGATCCAT 238
F6_020425       AAAAAAAACGGGGCTGGATGATACGAAAACGGTCTGAGACGGACTTATGTCCAGATCCAT 236
                *************** ************************************************

HHV7_RK_U69     CTGTAACAGACCTCCTATGTCATGAATCTTTGACTGTATCTCCAAAGTTTGAACGAGA-T 297
F6_020425       CTGTAACAGACCTCCTATGTCATGAATCTTTGACTGTATCTCCAAAGTTTGAACGANAAT 296
                ************************************************************ * *

HHV7_RK_U69     GGATTGAGTGCATGCACGGAATTTGAGAATTTTATGGATACAAGGAAAATCGTGTTAAGT 357
F6_020425       GGATTGAGTGCATGA------------------------------------------- 311
                ***************
```

HHV7U70

```
HHV7_RK_U70     ATGGCAATAGATTACGCACAAATTTCTTGTAATTTGGCTT---CTATTATAGAAGAGGAC 57
F7_020425       TTGGCAATAGATTACGCACAAATTTCTTGTAATTTGGNNAATANTATTATAGAAGAGGAC 60
                 ************************************       ****************

HHV7_RK_U70     TCGGTCTTTTTATTTCTAATAGACAAATTAAACAATCTGGACATTTCAAGAAGGAAAATT 117
F7_020425       TCGGTCTTTTTATTTCTAATAGACAAATTAAACAATCTGGACATTTCAAGAAGGAAAATT 120
                ************************************************************

HHV7_RK_U70     TCATTTAATTTTATCCGTCTGTGTTATACTTATTATATCTTAATAAAGTTTAATTCTCGC 177
F7_020425       TCATTTAATTTTATCCGTCTGTGTTATACTTATTATATCTTAATAAAGTTTAATTCTCGC 180
                ************************************************************

HHV7_RK_U70     TTCAAAGATACCTTCTTAGCCAGATCGTTCATTGATTATATGCATCAAAACATATCAGAT 237
F7_020425       TTCAAAGATACCTTCTTAGCCAGATCGTTCATTGATTATATGCATCAAAACATATCAGAT 240
                ************************************************************

HHV7_RK_U70     TTTATCGATGAGAATGTTGAGCTATCTGATTTATATAGCAATATTTATGT-CCGCTTACA 296
F7_020425       TTTATCGATGAGAATGTTGAGCTATCTGATTTATATAGCAATATTTATGTACCGCTTACC 300
```

```
                  ************************************************** ********

HHV7_RK_U70       AGATGCGAGTCCAAAAGTTGTTAAGAATCTATTTAAAATATTAGAACGAGAGACAAGAGG 356
F7_020425         AGATGCGA-------------------------------------------------- 308
                  ********

HHV7U71

HHV7_RK_U71       -ATGGGCTCAAAGTGCTGTAAGACAATACATGGCGGCATCTTCAGCAAGGCCGAAGACAC 59
F8_020425         TATGGGCTCAAAGTGCTGTAAGACAATACATGGCGGCATCTTCAGCAAGGCCGAAGACAC 60
                   ************************************************************

HHV7_RK_U71       CTTAGTAGACTATAAAGGAAAATATATTAATCTTGAAAAAGAATTTTCTGCTTTAAGTGA 119
F8_020425         CTTAGTAGACTATAAAGGAAAATATATTAATCTTGAAAAAGAATTTTCTGCTTTAAGTGA 120
                  ************************************************************

HHV7_RK_U71       TACTGAATCTGAAGAAGAGTTGCAACTAGAGAAGCCACTTCTAAATAAACAAGATTCTAG 179
F8_020425         TACTGAATCTGAAGAAGAGTTGCAACTAGAGAAGCCACTTCTAAATAAACAAGATTCTAG 180
                  ************************************************************

HHV7_RK_U71       CGTTTCGTTAACCCAGAAGAAACTTGAAAATCAATCCAAATAA 222
F8_020425         CGTTTCGTTAACCCAGAAGAAA---------------------- 202
                  **********************

HHV7U72

HHV7_RK_U72       ACACGATAAACTATCCCCCATCCAAGTTATATCCCTTGTAGATTGATTGATGTTGGGAGT 720
F9_020502         ------------TTCCCCCATCCAAGTTATATCCCTTGTAGATTGATTGATGTTGGGAGT 48
                             ************************************************

HHV7_RK_U72       TTTTTCGGAAGAGAATACTTTTACGCAACAAACTATATAGTATACTAAGATGAAGAAAAA 780
F9_020502         TTTTTCGGAAGAGAATACTTTTACGCAACAAACTATATAGTATACTAAGATGAAGAAAAA 108
                  ************************************************************

HHV7_RK_U72       GCAGATATCCATGAAAACAACATAGCAGATAAGTTGGATAGGATTTAGGTAGAGTTGTGG 840
F9_020502         GCAGATATCCATGAAAACAACATAGCAGATAAGTTGGATAGGATTTAGGTAGAGTTGTGG 168
                  ************************************************************

HHV7_RK_U72       CGTTAGGCTGCGGATGTCATTCCGTATAGAGAGGTTGATGGCTTTCATATCATTGATTTC 900
F9_020502         CGTTAGGCTGCGGATGTCATTCCGTATAGAGAGGTTGATGGCTTTCATATCATTGATTTC 228
                  ************************************************************

HHV7_RK_U72       ATAATATGCGCATGGGAATCCCAAATTGGGAAAGTGAACAGCCACTAAATGCACTGTGAC 960
F9_020502         ATAATATGCGCATGGGAATCCCAAATTGGGAAAGTGAACAGCCACTAAATGCACTGTGAC 288
                  ************************************************************

HHV7_RK_U72       GTTGACATACGTTAAACAAGCACAGATTATACTTAAGACCCAAATTCTCATATTAATCAC 1020
F9_020502         GTTGACATACGTTAAACAAGCACAGATTATACTTAAGACCCAAATTCTCATATTAATCAC 348
                  ************************************************************

HHV7_RK_U72       ATCTACTCTGCTCAATGCCAT 1041
F9_020502         ATCTACTCTGCTCAATGCCAA 369
                  **********************

HHV7U73

HHV7_RK_U73       ATGGAAACTCAGCTTCAAAATGATCAATTATTTTTGGAATGGTTTGGACAAAATTTGTTG 60
F10_020502        ----------------------------TTTTTTGGAATGGTTTGGACAAAATTTGTTG 31
                                              *******************************

HHV7_RK_U73       GATTGCCACTTTGCACAAAACGTTTCAGTGTATTTGCAAGATGCTTCAATGGTTCATTTT 120
F10_020502        GATTGCCACTTTGCACAAAACGTTTCAGTGTATTTGCAAGATGCTTCAATGGTTCATTTT 91
                  ************************************************************

HHV7_RK_U73       AAAACCTTTTCTGAACAAATAAAAATTATAAGAGCTCCAATGGGTTCTGGCAAAACCTCT 180
F10_020502        AAAACCTTTTCTGAACAAATAAAAATTATAAGAGCTCCAATGGGTTCTGGCAAAACCTCT 151
                  ************************************************************

HHV7_RK_U73       GCATTGATAGAATTCTTGAAAACTGTTTCATATATTGATTCTGTCCTTGTTATTTCCTGT 240
F10_020502        GCATTGATAGAATTCTTGAAAACTGTTTCATATATTGATTCTGTCCTTGTTATTTCCTGT 211
                  ************************************************************

HHV7_RK_U73       CGTAAAACTTTTGCTGCAGAACTTTTAAATAGATTCAAGAAGAATGATTTGAACGATTTC 300
F10_020502        CGTAAAACTTTTGCTGCAGAACTTTTAAATAGATTCAAGAAGAATGATTTGAACGATTTC 271
                  ************************************************************

HHV7_RK_U73       TATCTGTACAGCGAAATTAAAGAGCGTCAAATCAACAAGAACAAACTGATAATTCAAGTA 360
F10_020502        TATCTGTACAGCGAAATTAAAGAGCGTCA------------------------------ 300
                  ********************************

HHV7U74

HHV7_RK_U74       AATAAACAATATAGAGTACCAATTTATCAATGCCTGTTTTTTAATG--CTGAGACT-CAC 117
F11_020314        ----------------------------------------TNATGGCTTGAGACTTCAC 19
                                                        * ***   ******* ***

HHV7_RK_U74       TCCTTGCATACTTTT----TTGGTTATTGGAAATGAGATTTCTGAAAATCTCCTTGATAA 173
F11_020314        TCCTTGCATACNNATAAAATNGGTTATTGGAAATGAGATTTCTGAAAATCTCCTTGATAA 79
                  ***********   *    * ****************************************

HHV7_RK_U74       GATTTCCATTTCAAAAGAAAAAA-TCTTTATGTGGGATTTGAATGAACAAATAATGAGTA 232
F11_020314        GATTTCCATTTCAAAAGAAAAAAATCTTTATGTGGGATTTGAATGAACAAATAATGAGTA 139
                  ********************** ************************************

HHV7_RK_U74       TAACTCAAAAGACAGAGTCAATCTGTGAACTGATGTTTGGCAACAAACAAATAATTCAAA 292
F11_020314        TAACTCAAAAGACAGAGTCAATCTGTGAACTGATGTTTGGCAACAAACAAATAATTCAAA 199
                  ************************************************************
```

```
HHV7_RK_U74     AACTATCCAGAACATTTATTTTCTTAACAATTATTTTAAATTCTAATATGTATGAAATAA 352
F11_020314      AACTATCCAGAACATTTATTTTCNTA-CAATTATTTTAA-TTCTA-TATGTATGAA-TAT 255
                ********************** ** ************ ***** ********** **

HHV7_RK_U74     TAGATGTCTGCATAGACAGTAATGCGGTTTTGTACATGCCAGACTTAAGACCAATGATAA 412
F11_020314      AANTTGCATAACGTATCANNGAC----------------------------------- 278
                 *  **  *       **   *
```

HHV7U75

```
HHV7_RK_U75     ACTCAGCCGAATCTCTTCTAGCGTATCTTCGGAATCAAGAAGAATAATGCCAATGAAGCC 540
F12_020314      TCTCAGCCGAATCTCTTCTAGCGTATCTTCGGAATCAAGAAGAATAATGCCAATGAAGCC 60
                 ************************************************************

HHV7_RK_U75     TAGAAACCCTAATGTTTTCTCACATATGCTTTTATAGGATTTGCCATTTACAATTGTGCC 600
F12_020314      TAGAAACCCTAATGTTTTCTCACATATGCTTTTATAGGATTGCCGATTTACAATTGTGCC 120
                ******************************************  * ***************

HHV7_RK_U75     ATCTTCCAAATGAAACCATACATCATCATTTCCAACCGAAA--TATCAATTAA-AGGTAA 657
F12_020314      ATCTTCCAAATGAAACCATACATCATCATTCTCAACCGAAAACTATNATNNGATNNNTNN 180
                ********************************  *********  *** *    *    *

HHV7_RK_U75     ATTTTTAATGGAAACGTGTTCGGCGTGTTTGTTTAGTTCACAGATATTCCGGACGCTTGA 717
F12_020314      ANGGTTATGGGNGNAGTGC-CGGNG---------------------------------- 204
                *   ***  **    ***  *** *
```

HHV7U76

```
HHV7_RK_U76     AATCAATCTAGAGGAAATACTTTTTGCTAAGGGTGTCAGCGTA-ACCTTGTTTTTTGGTA 1379
G1_020228       --------------------------------GTGNCGGNNTTTGGCATACCGTNNTGGN 28
                                               *** * *  *    * *    *   *

HHV7_RK_U76     ATGTGACAATTGGTGTTATTCCAATCGTACAAATCCAATCAATGTATTTATTATAAATGT 1439
G1_020228       ANNNGACA--TGGTGT-ATTCCA-TCGTACAAATCCAATCAATGTATTTATTATAAATGT 84
                *   ****  ****** ****** ************************************

HHV7_RK_U76     CGTTATCGAGATT-TTGAAGGATGTAGTTTATCGTTTGCAAAATGCCTTCCAGGATTACC 1498
G1_020228       CGTTATCGAGATTCTTGAAGGATGTAGTTTATCGTTTGCAAAATGCCTTCCAGGATTACC 144
                ************* ************************************************

HHV7_RK_U76     TGTTTGATTGCATAATACCAGATATTAAAATATTCTTGAAATGACATGCCTTTACTGCGG 1558
G1_020228       TGTTTGATTGCATAATACCAGATATTAAAATATTCTTGAAATGACATGCCTTTACTGCGG 204
                ************************************************************

HHV7_RK_U76     AACGGAAATGCTTTCAGATGACGTTGCCATTCAGCATTGATCTGCTGGATGGATATACTA 1618
G1_020228       AACGGAAATGCTTTCAGATGACGTTGCCATTCAGCATTGATCTGCTGGATGGATATACTA 264
                ************************************************************

HHV7_RK_U76     TTCAAGAGAGTGTGTGACAGTGTGTAAAACATTTGCTTCTTAAAGACAGCCGAGTCTCTA 1678
G1_020228       TTCAAGAGAGTGTGTGACAGTGTGTAAAACATTTGCTTCTTAAAGACAGCCGAGTCTCTA 324
                ************************************************************

HHV7_RK_U76     AGAACACTGTGTATAGATTGTCCCTTAAAGAATTCTTTTTTACCGTGTAGCATTTCTATA 1738
G1_020228       AGAACACTGTGTATAGATTGTCCCTA---------------------------------- 350
                *************************
```

HHV7U77

```
HHV7_RK_U77     TTACCTTTTAGAGTATTATTAATAACCGGTACTGCGGGTGCGGGCAAAACAAGCAGCATT 300
G2_020425       ----------------------------------------TGCNAANACNAGCAGCATT 19
                                                        *  ** ** *********

HHV7_RK_U77     CAAACCTTAGCAGCTAATAGTGATTGTCTTATAACTGCTACCACTTCCATTGCTGCTCAA 360
G2_020425       CAAACCTTAGCAGCTAATAGTGATTGTCTTATAACTGCTACCACTTCCATTGCTGCTCAA 79
                ************************************************************

HHV7_RK_U77     AATCTTAGTGGTTTATTGAACAGAACCAAATCTGCGCAAGTGAAAACAATTTTTAAAACA 420
G2_020425       AATCTTAGTGGTTTATTGAACAGAACCAAATCTGCGCAAGTGAAAACAATTTTTAAAACA 139
                ************************************************************

HHV7_RK_U77     TTTGGTTTTAATAGTTCACATGTATCTATGAATGAAAGAATTAGTTGTTCAGTAACAACT 480
G2_020425       TTTGGTTTTAATAGTTCACATGTATCTATGAATGAAAGAATTAGTTGTTCAGTAACAACT 199
                ************************************************************

HHV7_RK_U77     TTAGATTCGATTGCGGATCAGCAAAAACATGATTTATCTACATATTGGAACGTCATCGCA 540
G2_020425       TTAGATTCGATTGCGGATCAGCAAAAACATGATTTATCTACATATTGGAACGTCATCGCA 259
                ************************************************************

HHV7_RK_U77     GATATAGCGGAAAGAGCTTTGAATGCAGCAAATGGGAAAGCCAAAGTGATACCTGATCTA 600
G2_020425       GATATAGCGGAAAGAGCTTTGAATGCGGCAAATGGGAAAGCCAAAGTGA----------- 308
                *************************** **********************
```

HHV7U79

```
HHV7_RK_U79     ATGATTGCGGAGGATAGAGAATATGGAACGTTTGAATC-TGTAACCCAGGCTTATCAGCA 59
G3_020313       GNAAGCANGANTGCGGGAGGATAGANANAGGNAGAAGNATCGAACCCAGNGNNANCAGCA 60
                   *    *   *   *** ***   *  *   ***   *  *******    * *****

HHV7_RK_U79     GATCATTAGTCATACTTTACAGCTGAGACGATATGAATTTGAAACTGGGTGCATGATTAT 119
G3_020313       GANNNNANANGTCNNCNACANGCGGAGACGANATNNANN--AAACGGGGCAGATAGNNN- 117
                **                    ** ******* **  *    **** ***   **

HHV7_RK_U79     GTTTTCTGCTAATTCTGGAAAATGTGAGATGCTCTCTAATGGTTGGATTTCAATGATTTC 179
G3_020313       -----CCGNNAATC---GGAAAAGGGAGA-GCTCTCTAA-GGNNGGATTCAA--GANNGC 165
                     * *  ***    * *** * **** ********* **  *****  *  **   *

HHV7_RK_U79     ATGGACTTCAGAAACCGATACGGCCGGCTCATTGACATTGGATATTTGTACTGAGGGAGG 239
```

```
G3_020313       ANGGACT-CAGAAACCGACACGGCCGGCTCAGGGNCAN-GNANCCCNACNCGAGGGGNNG 223
                * ****** ********** ************  * **  * *          *   ***  *

HHV7_RK_U79     GCAGTGCAAAACTTACAGTGCCAGAGGTCATATATTATGTTCTAAAAACATCACTTCAAT 299
G3_020313       GCAACACGGC-CGGGCAAAANNACC--CCATCCCCCACCCCCNCCCCANNCANCNTCCNN 280
                ***   *    *   **     *     ***     *    *      *     * **

HHV7_RK_U79     TTCTCAGAAAAACGAGGGAAAGGAAAAAGTTTTGACGATTTGTCATGACAATGGAAAATT 359
G3_020313       NTTTCTGNTTTACNTTTNTT---------------------------------------- 300
                 * ** *    **

HHV7U81

HHV7_RK_U81     TAAAGAATTCGGAGGGGAACATCCTCTCACAGTGCTGAAAGCTAATCCGTGACCTCTACC 480
G4_020425       ------------------------------TTGCTGAAAGCTAATCCGTGACCTCTACC 29
                                              ****************************

HHV7_RK_U81     ATCGGGATAAGGATCTTGGCCCACTATTATCACCTTAATCTCTTCAGGCGAACATAAATA 540
G4_020425       ATCGGGATAAGGATCTTGGCCCACTATTATCACCTTAATCTCTTCAGGCGAACATAAATA 89
                ************************************************************

HHV7_RK_U81     ACTCCAGCTGTGTACATTCTGAGGATCGGGGTAGATAATTAATCTTTCTCTGTCACGCTG 600
G4_020425       ACTCCAGCTGTGTACATTCTGAGGATCGGGGTAGATAATTAATCTTTCTCTGTCACGCTG 149
                ************************************************************

HHV7_RK_U81     AACCAATTTGTATACATTTTGTAATTGTACAATGTCAGAATCTGATAAGTTTAAAAACTT 660
G4_020425       AACCAATTTGTATACATTTTGTAATTGTACAATGTCAGAATCTGATAAGTTTAAAAACTT 209
                ************************************************************

HHV7_RK_U81     AAGCCACTTCACATTGATTTGAAACGTTTCATGTTGTTCTTCAAGAGACATTTTTATTGA 720
G4_020425       AAGCCACTTCACATTGATTTGAAACGTTTCATGTTGTTCTTCAAGAGACATTTTTATTGA 269
                ************************************************************

HHV7_RK_U81     GTTGTCATCTGAAATGTTTTCCAGCATCCACTGTAGTAATGCCAT 765
G4_020425       GTTGTCATCTGAAATGTTTTCCAGCATCCA--------------- 299
                *********************************

HHV7U82

HHV7_RK_U82     GAGTTTTTCATAAGAGTATGATTTACACACATTATCTCTACAGGTTAAAAGTGTGTTATC 360
G5_020419       ----------------------------------------------GGGANAATNNNNTTTTT 17
                                                               *  * **     ** *

HHV7_RK_U82     TGAAAGACATCTTTCAAAGTTATTTAAAAAGCCAAG------CCAGTTGTGCCCTGATTC 414
G5_020419       TAAAA----TTTTTAAAGGGGAAANAAAGGTTTAAAGGGGGTTGGGTTGGGTAAGGNTTG 73
                * ***    * *** ** * *   ***     **          **** *    * **

HHV7_RK_U82     TGAATCTTTAATTAAGGAAAGTAAGGTGTGTAGTTGATTAGGATTATTATACAGTAGGTA 474
G5_020419       GGAAATNTTAAG-GGGAANNTTAANANTAANAACCGANNTAAGGGGTTANN-AATGGTTN 131
                 ***   ****    * *   ***       *   **          ***   * * * *

HHV7_RK_U82     AAATGTTTCTAAAAAATCTTCATCAA--TTAAAACTT-GATCTCTCCTTGAGATATAGGG 531
G5_020419       --TTCTTTNNAANNTCTTTTCATCGNAGTTAAAACTNTGATCTCTCCTTGAGATATAGGG 189
                   * ***  **    * ******    ********  **********************

HHV7_RK_U82     AG-TATGATATAATAAGCGAGAAATTATATCCGAATTTGGTATAGGTGCCTGATCAA--T 588
G5_020419       ANGTATGATATAATAAGCGAGAAATTATATCCGAATTTGGTATAGGTGCCTGATCGAATN 249
                *  ************************************************** *

HHV7_RK_U82     AGGTTCCAGAAGCGACAGAGTTTTTGATTGGGTACAATTTAG-AATGTTTTTGATACATT 647
G5_020419       NGGTTCCAGAAGCGACGGAGTTTTTGATTGGGTACAATTTAGGAATGTTTTTGATACATT 309
                 *************** ************************* *****************

HHV7_RK_U82     CTTGTTCTAAATTAG-AATAATAAGAG--TTATTAAA-TAAAG-CATATATTTGGTTT-- 700
G5_020419       CTTGTTCTAAATTAGTAATAATAAGTAGTTTATTANAATAAAGTCATATATTTGGTTTTN 369
                *************** *********    ****** * ***** **************

HHV7_RK_U82     ----AGAATGGAGATAAGAAATATAAAAAA-GATGTTAGTTTTCAT-------------- 741
G5_020419       NTAANNNANGNTGATAGATAANATAATCACTGANGTTAATTGCNGCGTNTCGCTCCGTNA 429
                       * *  ****   ** ****  *  ** **** **

HHV7_RK_U82     -------
G5_020419       CGCGTTT 436

HHV7U84

HHV7_RK_U84     GTTAAAATTGCTGTGAACAATACCCTTAGATTTTGTTGCAAAGTTTGT--CGTTGAGAGA 658
G6_020502       ------------------------------TTTGTTGCAAAGTTTGNTACGTTGAGAGA 29
                                              *****************   **********

HHV7_RK_U84     TTTTCTTT--TACAGCCTTAAATATTTTTCTCAATTTAGAATGTTTATAGATGGCATTCG 716
G6_020502       TTTTGNNNGTTACAGCCTTAAATATTTTTCTCAATTTAGAATGTTTATAGATGGCATTCG 89
                ****       **************************************************

HHV7_RK_U84     TATATTTGTCCGCCTTAAGTAACATTTCTGATTTTGAAATGTTTATTTCAGAATTTATTT 776
G6_020502       TATATTTGTCCGCCTTAAGTAACATTTCTGATTTTGAAATGTTTATTTCAGAATTTATTT 149
                ************************************************************

HHV7_RK_U84     TTTCAGGAATTGTACATTTTATGGCTTTTTCATTAAACGGTAACGTTTTTGAAAAACAGT 836
G6_020502       TTTCAGGAATTGTACATTTTATGGCTTTTTCATTAAACGGTAACGTTTTTGAAAAACAGT 209
                ************************************************************

HHV7_RK_U84     TACATGTATATCGAGCATTGGCTTTAGTACACATGCAGCTTTTAGAAGTGAGTTTACCTT 896
G6_020502       TACATGTATATCGAGCATTGGCTTTAGTACACATGCAGCTTTTAGAAGTGAGTTTANCCT 269
                ************************************************************ * *

HHV7_RK_U84     T---GGCCTTACAAACAGTTTTACGTTTTTTGCCCCTCAT 933
G6_020502       TTGGANACTTACAAACAGTTA------------------- 290
                *      **************
```

HHV7U85

```
HHV7_RK_U85     AGTCTTGGCGAAGAATGCACATGTTAAGCAACCAGCATCGTTAATTGTTCGAGATTTAAA 540
G7_020419       ----------------------------TCCAGCATCGTTAATTGTTCGAGATTTAAA 30
                                            ****************************

HHV7_RK_U85     TTTCAGCAATGCTGTGATATTGTTAGCATTGAAGTAAATAATGCT-TTCTTTGTCACGAT 599
G7_020419       TTTCAGCAATGCTGTGATATTGTTAGCATTGAAGTAAATAATGCTCTTCTTTGTCACGAT 90
                ********************************************* **************

HHV7_RK_U85     GTTTAAATTCATGAACTATTTTTTCACTCATGTTTTTCA-GGTTTTTGGTATTTATGGTA 658
G7_020419       GTTTAAATTCATGAACTATTTTTTCACTCATGTTTTTCACGGTTTTTGGTATTTATGGTA 150
                ****************************************** *********************

HHV7_RK_U85     TGATTTAATATCAGTTCGTAATGTTTAAACCATATTACAGAGCTGAGATT-TTTAGTTGT 717
G7_020419       TGATTTAATATCAGTTCGTAATGTTTAAACCATATTACAGAGCTGAGATTCTTTAGTTGT 210
                ************************************************** *********

HHV7_RK_U85     ATCAACAGTACACTGAAATACAATTGTGTTTCCTAGTTCTACTCTATGATGCATTATTGG 777
G7_020419       ATCAACAGTACACTGAAATACAATTGTGTTTCCTAGTTCTACTCTATGATGCATTATTGG 270
                ************************************************************

HHV7_RK_U85     AGATAATGCATATATGATTACAGAAAATCCAATCATCTTTATGAAAATTGCACCTAAGAA 837
G7_020419       ATATAATGCATATATGATTACAGAAAATCCAATCATCTTTATGAAAATTGCACCTAAGAA 330
                * ************************************************************

HHV7_RK_U85     AGGCAT 843
G7_020419       AGGCAA 336
                ******
```

HHV7U86

```
HHV7_RK_U86     TCTGTTGTTGCGACCCCTGTAATGATGTTGTCGATAGGTTGTATCGAACGAATTGTGACA 3240
G8_020314       ------------------------------------------------------CCCNANA 7
                                                                       * *

HHV7_RK_U86     AAGTTCTCTATTTTTATAGTCCAAATTGTTAAAAATATCAACTGAAGATTGGGGTTGATA 3300
G8_020314       AAAAAGAGNANNNNAAAA----AATTGGAAACAAAAGAAAAAANAANANGAGGGATAAGA 63
                **       *     * *    ** * *  * ***    **   ** *   *** * * *

HHV7_RK_U86     TGGTAAAGAGATAAAATTTTCTCCACTTGCTTGAGGATAATTATAATACAATGATTGCAG 3360
G8_020314       TGATAAAAAAAAAAAA-----------GGAGGTNGAGGAANANANTATANTGATG---- 107
                ** **** * * ****            *   *  **  *  * * ** * ****

HHV7_RK_U86     TCCTTGATGTGGCAATTCTGATGCAAATCGAAATTGTCCAGTTTCAGTTGTATTTGGAGA 3420
G8_020314       -CGTCGATGTGGCA--TNNGATGCA--TCNAA--TGTCGTNNNNTNGNNNNNNT-GGAGA 159
                 * * *********  *  ******  ** **  ****       *      * *****

HHV7_RK_U86     AGATTCAGAAGTAAAAGTACTTCTTGATTTAAGAGGAATAGCACGTCGCGGATCTTGAA- 3479
G8_020314       GATCCCCCGNNTNNT----CTGGTTGATTTANGAGGAATAGCACGTCGCGGATCTTGANC 215
                     *    *       **  ******** **************************

HHV7_RK_U86     -GAGTTTTTGATGTGACTATTTC-TTGGTTGTTAAAATTTGAATAAGTTATTTCATTTTC 3537
G8_020314       NANGTTTTTGATGTGACTATTTCCTTGGTTGTTN---TTTGACTCCGTNTTTCATTTTCN 272
                   ******************** *********    ***** *  **  **   ***

HHV7_RK_U86     TGGTTTTGTATGATC-GTGAATTTC----ATTATATAAACCCGAAAGA-GTCATGGCTTC 3591
G8_020314       TGGTTTTGTNTGATCCGTGNAATCCCCCCATAATATNNCCTTTTTTTTTAGTNTGGCTTC 332
                ********* ***** *** * * *    ** ****   *             *******

HHV7_RK_U86     ACGAACCACGTCAGCTGTAGTGCTCAT---------------------------------- 3618
G8_020314       CGANCNCGTAGCNGTNTGCATNTTCNNCNGNNNTGGNGCGGNCTNNCNGTNNGGCCNNTA 392
                      *   * *      *  **
```

HHV7U90

```
HHV7_RK_U90     AGCTAGAATTTAAAGAAAAAAGTCTAAATATGAAATCTGTACACATATATCATATTTTAT 3360
G9_020502       ------------------------------------------------------CGTCCT-C 7
                                                                       * *

HHV7_RK_U90     ATGTAAGAAAATGAAAAAAAAATACCTGCTGCCAGTTTAATATCCGAGGAAGCTCCTTCT 3420
G9_020502       AGGNCNCANCGCNCGGGCGNCNCGCATATNCTCA---TNNNTCCANGAANNGCCNTTNCT 64
                * *    *                 * *     **   *     *       **   * **

HHV7_RK_U90     TCAACATTTGTGGAATAGGAAGCATTTGCCGTTGCAATATCATTAGATGATATTGAGTCA 3480
G9_020502       TCAACATTTGTGGAATAGGAAGCATTTGCCGTTGCAATATCATTAGATGATATTGAGTCA 124
                ************************************************************

HHV7_RK_U90     AGGAAATCAAAAGAATCCAATTCTTGGAAGAGAGTATTATCAAAGGAAGGGTTGCATTTT 3540
G9_020502       AGGAAATCAAAAGAATCCAATTCTTGGAAGAGAGTATTATCAAAGGAAGGGTTGCATTTT 184
                ************************************************************

HHV7_RK_U90     AAACTGTCTTGCAATATGTCTCTGATGGTCTCTTCCACGGTGGTATCTGGGTGATATCTG 3600
G9_020502       AAACTGTCTTGCAATATGTCTCTGATGGTCTCTTCCACGGTGGTATCTGGGTGATATCTG 244
                ************************************************************

HHV7_RK_U90     TAATAAAATTTATAAATTTTATTACAGAGAAATTAAATCTGTCTTTTTTGTATTAAAATA 3660
G9_020502       TAATAAAATTTATAAATTTTATTACAGAGAAATTAAATCTGTCTTTTTTGTATTAAAATA 304
                ************************************************************

HHV7_RK_U90     TTTAGGTTTACCTACTCAATATGTCTTTCGTCATTCGATGCCATGGTTTCAGATGGTTGT 3720
G9_020502       TTTAGNTTTACCTACTCAATATGTCTTTCGTCATTCGATGCCATGGTTTCAGATGGTTGT 364
                ***** ******************************************************

HHV7_RK_U90     CCAACAATCAACATAGGTGAAGTAGCTCCACTTCTTTCCAT 3761
G9_020502       CCAA---------------------------------- 368
                ****
```

HHV7U91

```
HHV7_RK_U91     ATGTATACTCTGGAATATGAAAAACGTGTATCAAGGCCAAAACTTACTTATTGGATCATT 60
G10_020425      --------------------TAAACGTGTATNNAGGCCAAAACTTACTTATTGGATCATT 40
                                    **********  ****************************

HHV7_RK_U91     TTGGCAATATTATTTGTTTTTTTAATAATAACTGGATCTGTATTAATTGTAATTGAAACT 120
G10_020425      TTGGCAATATTATTTGTTTTTTTAATAATAACTGGATCTGTATTAATTGTAATTGAAACT 100
                ************************************************************

HHV7_RK_U91     CTGTCAATTCAAAGAACGACTTTGAATGCTCAAAATGATAAAACAAGTACTGTTGTTCCA 180
G10_020425      CTGTCAATTCAAAGAACGACTTTGAATGCTCAAAATGATAAAACAAGTACTGTTGTTCCA 160
                ************************************************************

HHV7_RK_U91     GAATTAACATCTAATTCTCCAGGTTTGTTCAGCTATAAAATGTCATTGTTTAAAAAAAAA 240
G10_020425      GAATTAACATCTAATTCTCCAGGTTTGTTCAGCTATAAAATGTCATTGTTTAAAAAAAAA 220
                ************************************************************

HHV7_RK_U91     GTTTGTAACTCTGTTAATATGAATATTTTTTCTTGTAGATCAAACAACTGTGACAAATTT 300
G10_020425      GTTTGTAACTCTGTTAATATGAATATTTTTTCTTGTAGATCAAACAACTGTGACAAATTT 280
                ************************************************************

HHV7_RK_U91     TTCTGCAAGTTCTAAACCAACTCTTAGCAGTAAACA-ACCCGGATGGATACAGGCACTAA 359
G10_020425      TTCTGCAAGTTCTAAACCAACTCTTAGCAGTAAACACACCCGGATGA------------- 327
                ************************************ *********
```

HHV7U95

```
HHV7_RK_U95     --ATGGAGAGTGGGGATAGTTTCGGAAATAATCA---CCAAGTCTCATCGAATT-CAGAT 54
G11_020314      ATGGGAGGAGTGGGGATAGNNTCGGAAATAAAANAANCCAAGTCTCATCGAATTTCAGAN 60
                    *  ************  **********       ****************** ****

HHV7_RK_U95     GCTTTTCAATTCCGGCAATATTCAACAACTGTAGATACTTTTGCATATTCATCAATGGAT 114
G11_020314      NCTNTCNAATTCCGGCAATATTCAACAACTGTAGATACN---GCATATTCATCAATGGAT 117
                 ** *  **********************************    ******************

HHV7_RK_U95     CC--TTCAAATTGCATGCT--------TAATGAACA-AACTCATTTAGAAACGGTACTAG 163
G11_020314      NNGCTTCAAATTGTCTATGGGGGGGGATGNTTATCATAACTCATTTAGAAACGGTACTAG 177
                    *********  *            *  * * ** ************************

HHV7_RK_U95     TTTATCCTGTTTACAACATG-TCT-TCCCAGCA-AACAGTGGAGAGAAATCTCTATGCCA 220
G11_020314      TTTATCCTGTGTACAACAGGGTCTGTCCCAGCCCAACAGTGGAGAGAAATCTCTATGCCA 237
                ********** ******* * *** *******  ****************************

HHV7_RK_U95     ATAATATAGGAAATGGAAATGTGCTGAGTTATTGTAGACAAGATAGTTTGCAAGGAAACC 280
G11_020314      ATAATATAGGAAATGGAAATGTGC-GAGTTATTGTAGANCAGATAGTGNGGGGGGAAGGC 296
                ************************ *************  *******  *   ****  *

HHV7_RK_U95     --AAGGAACAT-ATTTTCCAACATATCAGGGATTTCACTCAAACACTCCTCATAACTTTT 337
G11_020314      GCNNGNANCNTTATTATATGGTATNNTTNGNNTNTTTTTTTT-CTTTGNNNNTAATCCTT 355
                   * * * * *** *     **      *  * *   *     *  *      ***    **

HHV7_RK_U95     CTATAGAGAATTTTAATTTTGGCACCAATGTCATCAGGCCTATACCATTTAGAAATGAGG 397
G11_020314      TTNTT----TTTTTNNTTT------------------------------------------ 370
                 * *       ****  ***
```

HHV7U100

```
HHV7_RK_U100    AGAGAATAGTGTAAGCCTGTAGTTACTGACAAGACTTCAGACAGCCCTTTACTCATAATT 2280
G12_020314      ----------------------------------------------------CCCCCCCCT 9
                                                                    * *     *

HHV7_RK_U100    TTTTCTCTTAATCCGCGCACTCCCAATGAGTTCTCTTTTAGTGGAATGCAGAGTGTTGTT 2340
G12_020314      NCNCCCCTNCTNCCACCANCNCCAAATACANNAAANNAAAAAAAGANNAGNGNNGGGGNA 69
                    * **`   ** *   * ** ***           *     *        *  *

HHV7_RK_U100    TCTAATAATATAACCCTTAGAAGAATTAAAAGTAAGGTGTGAGAAATTTACTTAAAAAGC 2400
G12_020314      AAAANTANNAAAANAAANNNAANACAANAAAATN-GAAATGAGTGTNNNNTNCATAAANC 128
                  * **  * **       ** *    *** * *   ****         * *** *

HHV7_RK_U100    TCTTCAAAGTCGTCATATTGT-AGAAATCAGT---GTAACCATTTTTACT-TACTGTGCT 2455
G12_020314      NNNATCAGATNNNCATAGTGNTANAAATCNCCCCGNTAACCATTTTTACNCNACTGTGCT 188
                      *  *   **** **  * *****       *************   ********

HHV7_RK_U100    TCGTCTGGCCATTCACATTCGCAGGAAAACGATGTACATCTTGTGTATTCTAACAAGCTT 2515
G12_020314      TCGTCTGGCCATTCACATTCGCAGGAAAACGATGTACATCTTGTGTATTCTAACAAGCTT 248
                ************************************************************

HHV7_RK_U100    TGTCTTTCCTCGGAACTTAGACC-AATATATCTAGTAGAGCTTGCGTTGTGCATGTTTTC 2574
G12_020314      TGTCTTTCCTCGGAACTTAGACCCAATATATCTAGTANAGCTTGCGTTGTGCATGTCTTC 308
                ************************ ************* ****************** ***

HHV7_RK_U100    ATTTATT---AAGTTTTCGATTA-TATCTTCTAGAGTTATTGCCGTGCGTGCTTTTT-CA 2629
G12_020314      ATTTNTTTNNNNGTTTTCGATTAATATCTTCTAGAGTTATTGCCGTGCGTGCNNTTTTCA 368
                **** **     ************ ***************************  *** **

HHV7_RK_U100    AGTGGGACCGTATTGGTACATATCTTTATGAGTAATGCGAAAATTATATAATGCAGTTGC 2689
G12_020314      AGTGGGACCGTATTGGA------------------------------------------- 385
                ****************
```

Second attempt Results:

>CL7 = HHV7 C1 (U28)
>EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
Length = 153,080
Minus Strand HSPs:
Score = 1361 (210.3 bits), Expect = 3.0e-53, P = 3.0e-53
Identities = 273/274 (99%), Positives = 273/274 (99%), Strand = Minus / Plus

```
Query:   317 ATGGAATCAAATTCCGAAGCAGAGTTTGTTTTGAGATAGAACTGGCCTAACAAACGATCA 258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 40404 ATGGAATCAAATTCCGAAGCAGAGTTTGTTTTGAGATAGAACTGGCCTAACAAACGATCA 40463

Query:   257 ACGTGCGGTGCATCATGCATTGCTCCCAGTGCTGAAATTATGGCATCACGTAAGCACCAC 198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 40464 ACGTGCGGTGCATCATGCATTGCTCCCAGTGCTGAAATTATGGCATCACGTAAGCACCAC 40523

Query:   197 ATATCTAAGCCCGAATTTAAGCGCCGGAGAATTTTTGAAATTTCATATAACTGCTTATAA 138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 40524 ATATCTAAGCCCGAATTTAAGCGCCGGAGAATTTTTGAAATTTCATATAACTGCTTATAA 40583

Query:   137 AGAAGAGTTTTGTGATGGAAGCGTGAATAGCAAAACCGTTTTTTGTTATATTCATCATGA 78
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 40584 AGAAGAGTTTTGTGATGGAAGCGTGAATAGCAAAACCGTTTTTTGTTATATTCATCATGA 40643

Query:    77 CAAAGACATTTCCTCAAGAGACCCATGTTTTAAT 44
             ||||||||||||||||||||||||||||||| ||
Sbjct: 40644 CAAAGACATTTCCTCAAGAGACCCATGTTTTTAT 40677


>CL8 = HHV-7 C12 (U39)
EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 1422 (219.4 bits), Expect = 5.3e-56, P = 5.3e-56
 Identities = 290/297 (97%), Positives = 290/297 (97%), Strand = Plus / Plus

Query:    35 TTCACTAGTGATTTTGGCATTGGCAAAGTTGTAACAGTTCGATCTAAGAAATACACAGTG 94
             | ||||  |    ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 60735 TGCACTTCTTCAATTGGCATTGGCAAAGTTGTAACAGTTCGATCTAAGAAATACACAGTG 60794

Query:    95 CCAACATCGCGATATGTTGTTTGAAATGTAAGTTCTTTTTTAAAAGTTCTGACAGAAAAC 154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 60795 CCAACATCGCGATATGTTGTTTGAAATGTAAGTTCTTTTTTAAAAGTTCTGACAGAAAAC 60854

Query:   155 GTGTGTGCTTCAATCTTTGTTTTGTAAATTATCAAAATACCTTCGGTAGTTTTAATGTTA 214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 60855 GTGTGTGCTTCAATCTTTGTTTTGTAAATTATCAAAATACCTTCGGTAGTTTTAATGTTA 60914

Query:   215 GAGCCGTAAGACGCGCAAGAAACCTCCCTGTCAAAACGCACTAAATCTGTCCCGGTGGCA 274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 60915 GAGCCGTAAGACGCGCAAGAAACCTCCCTGTCAAAACGCACTAAATCTGTCCCGGTGGCA 60974

Query:   275 ATTGAACAAATTCGAAATGGTAAATGCTGATTGTGTCCAGTCATGACAAAGTCAGCT 331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 60975 ATTGAACAAATTCGAAATGGTAAATGCTGATTGTGTCCAGTCATGACAAAGTCAGCT 61031


>CL9 = HHV7 D4 (U43)
>EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Minus Strand HSPs:
Score = 1136 (176.5 bits), Expect = 4.4e-43, P = 4.4e-43
 Identities = 228/229 (99%), Positives = 228/229 (99%), Strand = Minus / Plus

Query:   229 AGTAAAAAATCTTTATACAGATTTTCTGTTTCTTCCGTATTTAAACAGTGAATAATATCC 170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71003 AGTAAAAAATCTTTATACAGATTTTCTGTTTCTTCCGTATTTAAACAGTGAATAATATCC 71062

Query:   169 TCAGACTTAATGGGTAATGCATTTTTTACATAATAGCTGAAATTCAAAGATTCTGCGTCG 110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71063 TCAGACTTAATGGGTAATGCATTTTTTACATAATAGCTGAAATTCAAAGATTCTGCGTCG 71122

Query:   109 CAAACAAATACCGTCTCAATTCTTTTGGAAAACTTGCATTTCTGTGTCTGCATACAAAAA 50
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71123 CAAACAAATACCGTCTCAATTCTTTTGGAAAACTTGCATTTCTGTGTCTGCATACAAAAA 71182

Query:    49 TATACATTTTTTGATGGTTTATATTTAACGATTATCGGAAAAATTAAAT 1
             |||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct: 71183 TATACATTTTTTGATGGTTTATATTTAACGATTATTGGAAAAATTAAAT 71231


>CL10 = HHV7 E8 (U57)

>EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Minus Strand HSPs:
Score = 542 (87.4 bits), Expect = 3.0e-16, P = 3.0e-16
 Identities = 110/112 (98%), Positives = 110/112 (98%), Strand = Minus / Plus

Query:   112 AAATCTATCCATTCAGTCCGATTACAGTAAATCCCAAGTAATGCTTCGAAACTTATGTTA 53
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 91585 AAATCTATCCATTCAGTCCGATTACAGTAAATCCCAAGTAATGCTTCGAAACTTATGTTA 91644

Query:    52 TAACGATCAGAGTTGTCGCCGTAATACAATCTTAAGTTTTCAAAAAGTTGTT 1
             |||||||||||||| ||| ||||||||||||||||||||||||||||||||
Sbjct: 91645 TAACGATCAGAGTCGTCACCGTAATACAATCTTAAGTTTTCAAAAAGTTGTT 91696
```

Appendix B – Identities of lymphochip clones on custom-made microarrays

Plate	Row	Column	Well	Clone ID	Name	GB Acc#
AP-hu01	A	1	A1	67037	Homo sapiens transcribed sequence with weak similarity to protein ref:NP_060265.1 (H.sapiens) hypothetical protein FLJ20378 [Homo sapiens]	T70413
AP-hu01	A	2	A2	246430	UDP glycosyltransferase 2 family, polypeptide B4	N73214
AP-hu01	A	3	A3	67071	interleukin 22 receptor, alpha 1	T70439
AP-hu01	A	4	A4	257298	0	N39874
AP-hu01	A	5	A5	110393	protein phosphatase 2 (formerly 2A), regulatory subunit B", alpha	T84132
AP-hu01	A	6	A6	257581	matrix metalloproteinase 12 (macrophage elastase)	N41372
AP-hu01	A	7	A7	111460	cyclin T1	T90767
AP-hu01	A	8	A8	257746	Homo sapiens transcribed sequences	N41840
AP-hu01	A	9	A9	111721	insulin induced gene 2	T91091
AP-hu01	A	10	A10	258790	cholecystokinin	N40841
AP-hu01	A	11	A11	112148	0	T92057
AP-hu01	A	12	A12	263341	DKFZP566K1924 protein	N20003
AP-hu01	A	13	A13	113786	hypothetical protein MGC45562	T77343
AP-hu01	A	14	A14	263836	zinc finger protein 24 (KOX 17)	N28492
AP-hu01	A	15	A15	114045	Homo sapiens transcribed sequences	T79505
AP-hu01	A	16	A16	264287	CCAAT-box-binding transcription factor	N27463
AP-hu01	A	17	A17	114048	erythropoietin receptor	T79507
AP-hu01	A	18	A18	264530	Pirin	N29255
AP-hu01	A	19	A19	114094	cytochrome P450, family 2, subfamily C, polypeptide 8	T79539
AP-hu01	A	20	A20	265553	TANK-binding kinase 1	N27804
AP-hu01	A	21	A21	114807	ATP-binding cassette, sub-family G (WHITE), member 2	T87360
AP-hu01	A	22	A22	266074	caldesmon 1	N31556
AP-hu01	A	23	A23	115614	guanine nucleotide binding protein (G protein), alpha 15 (Gq class)	T87268
AP-hu01	A	24	A24	268412	MCF.2 cell line derived transforming sequence	N35192
AP-hu01	B	1	B1	358379	associated molecule with the SH3 domain of STAM (AMSH) like protein	W95888
AP-hu01	B	2	B2	727307	thyroid hormone receptor interactor 13	AA401737
AP-hu01	B	3	B3	358433	retinoid X receptor, gamma	W96099
AP-hu01	B	4	B4	727705	lactotransferrin	AA398686
AP-hu01	B	5	B5	358506	collagen, type XVI, alpha 1	W96115
AP-hu01	B	6	B6	728707	Homo sapiens transcribed sequences	AA435905
AP-hu01	B	7	B7	359134	transforming, acidic coiled-coil containing protein 1	AA010120
AP-hu01	B	8	B8	729912	DNA (cytosine-5-)-methyltransferase 2	AA399591
AP-hu01	B	9	B9	359617	aldehyde dehydrogenase 9 family, member A1	AA010649
AP-hu01	B	10	B10	730346	heme binding protein 1	AA470092
AP-hu01	B	11	B11	359769	RAB interacting factor	AA010789
AP-hu01	B	12	B12	730397	TATA box binding protein	AA470112
AP-hu01	B	13	B13	360531	v-raf murine sarcoma 3611 viral oncogene homolog 1	AA015618
AP-hu01	B	14	B14	730950	chromosome 3 open reading frame 8	AA421118
AP-hu01	B	15	B15	360723	peroxiredoxin 6	AA016074
AP-hu01	B	16	B16	731225	chromosome 1 open reading frame 16	AA421062
AP-hu01	B	17	B17	362942	DIRAS family, GTP-binding RAS-like 2	AA019438
AP-hu01	B	18	B18	731592	protein phosphatase 1, regulatory (inhibitor) subunit 8	AA412525
AP-hu01	B	19	B19	364557	hypothetical protein MGC8721	AA022603
AP-hu01	B	20	B20	739891	E74-like factor 3 (ets domain transcription factor, epithelial-specific)	AA477878
AP-hu01	B	21	B21	365098	BCL2/adenovirus E1B 19kDa interacting protein 3-like	AA025195
AP-hu01	B	22	B22	740408	H2A histone family, member X	AA478223
AP-hu01	B	23	B23	365452	chromosome 3p21.1 gene sequence	AA009799
AP-hu01	B	24	B24	741406	Ras and Rab interactor 1	AA400958
AP-hu01	C	1	C1	116436	transcriptional adaptor 2 (ADA2 homolog, yeast)-like	T91418
AP-hu01	C	2	C2	269800	hypothetical protein FLJ10116	N36276
AP-hu01	C	3	C3	120386	0	T95837
AP-hu01	C	4	C4	270616	Rag D protein	N42899
AP-hu01	C	5	C5	121454	Homo sapiens transcribed sequence	AF143883
AP-hu01	C	6	C6	270939	catenin (cadherin-associated protein), beta 1, 88kDa	N42740
AP-hu01	C	7	C7	121948	interferon-related developmental regulator 1	T97868
AP-hu01	C	8	C8	271051	phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog, Drosophila)	N42967
AP-hu01	C	9	C9	122112	beta-transducin repeat containing	T98478
AP-hu01	C	10	C10	271985	tyrosinase (oculocutaneous albinism IA)	N42770
AP-hu01	C	11	C11	124086	interleukin 11 receptor, alpha	R02708
AP-hu01	C	12	C12	276501	hypothetical gene supported by AK026880; BC030095; AF052138	N43917
AP-hu01	C	13	C13	124708	vacuolar protein sorting 4B (yeast)	R02155
AP-hu01	C	14	C14	277058	hypothetical protein FLJ10781	N39582
AP-hu01	C	15	C15	126348	transmembrane 4 superfamily member 4	R06489
AP-hu01	C	16	C16	277138	protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), gamma isoform	N40919
AP-hu01	C	17	C17	127086	T-cell acute lymphocytic leukemia 1	R08058
AP-hu01	C	18	C18	277359	phosphoglycerate mutase 2 (muscle)	N57500
AP-hu01	C	19	C19	127841	pyruvate kinase, liver and RBC	R08829
AP-hu01	C	20	C20	278257	tumor protein p53 inducible nuclear protein 1	N94892
AP-hu01	C	21	C21	128065	BH3 interacting domain death agonist	R09650
AP-hu01	C	22	C22	278448	Homo sapiens transcribed sequences	N98717
AP-hu01	C	23	C23	128129	collagen, type II, alpha 1 (primary osteoarthritis, spondyloepiphyseal dysplasia, congenital)	R09767
AP-hu01	C	24	C24	279665	E3-binding protein	N49046
AP-hu01	D	1	D1	365837	G-rich RNA sequence binding factor 1	AA026052
AP-hu01	D	2	D2	742643	0	AA401504
AP-hu01	D	3	D3	365866	titin	AA025367
AP-hu01	D	4	D4	742747	capicua homolog (Drosophila)	AA400572
AP-hu01	D	5	D5	366785	chromosome 21 open reading frame 6	AA029407
AP-hu01	D	6	D6	743233	polymerase (DNA directed), delta 3	AA401552

AP-hu01	D	7	D7	376147	RAB5C, member RAS oncogene family	AA039472
AP-hu01	D	8	D8	743316	0	AA400583
AP-hu01	D	9	D9	376394	protein phosphatase 4 (formerly X), catalytic subunit	AA041409
AP-hu01	D	10	D10	752754	hypothetical protein LOC284611	AA482455
AP-hu01	D	11	D11	376405	chromosome 6 open reading frame 4	AA041420
AP-hu01	D	12	D12	753278	chemokine (C-X-C motif) ligand 16	AA411655
AP-hu01	D	13	D13	376522	gap junction protein, beta 1, 32kDa (connexin 32, Charcot-Marie-Tooth neuropathy, X-linked)	AA041502
AP-hu01	D	14	D14	753284	glutathione S-transferase M5	AA411662
AP-hu01	D	15	D15	376708	hypothetical protein BC013995	AA046358
AP-hu01	D	16	D16	753969	discs, large (Drosophila) homolog 5	AA478949
AP-hu01	D	17	D17	377282	stearoyl-CoA desaturase (delta-9-desaturase)	AA054956
AP-hu01	D	18	D18	754054	homeo box D9	AA478764
AP-hu01	D	19	D19	382913	v-ral simian leukemia viral oncogene homolog B (ras related; GTP binding protein)	AA084543
AP-hu01	D	20	D20	755663	retinoic acid receptor, beta	AA419238
AP-hu01	D	21	D21	383156	chloride channel 4	AA071201
AP-hu01	D	22	D22	756208	retinoic acid receptor, gamma	AA481863
AP-hu01	D	23	D23	415293	kininogen	W92148
AP-hu01	D	24	D24	756557	thiosulfate sulfurtransferase (rhodanese)	AA481682
AP-hu01	E	1	E1	128274	zinc finger protein, subfamily 1A, 5 (Pegasus)	R11517
AP-hu01	E	2	E2	280231	coagulation factor C homolog, cochlin (Limulus polyphemus)	N49197
AP-hu01	E	3	E3	128667	chemokine (C-C motif) ligand 14	R16807
AP-hu01	E	4	E4	280758	thymosin, beta 4, Y chromosome	N50611
AP-hu01	E	5	E5	128921	sterol O-acyltransferase 2	R10273
AP-hu01	E	6	E6	280868	Homo sapiens transcribed sequence with strong similarity to protein pir:S71101 (H.sapiens) S71101 protein kinase JNK2 isoform beta-1 - human	N47532
AP-hu01	E	7	E7	129585	Homo sapiens transcribed sequence with moderate similarity to protein sp:P02795 (H.sapiens) MT2_HUMAN Metallothionein-II (MT-II)	R16596
AP-hu01	E	8	E8	281043	Sapiens cDNA FLJ37676 fis, clone BRHIP2012627.	N50903
AP-hu01	E	9	E9	129900	0	R19169
AP-hu01	E	10	E10	281949	phosphoinositide-binding protein PIP3-E	N53325
AP-hu01	E	11	E11	130313	nitric oxide synthase 3 (endothelial cell)	R21248
AP-hu01	E	12	E12	284547	corticotropin releasing hormone binding protein	N76172
AP-hu01	E	13	E13	132660	aldehyde dehydrogenase 3 family, member B2	R25587
AP-hu01	E	14	E14	288654	general transcription factor IIE, polypeptide 2, beta 34kDa	N79351
AP-hu01	E	15	E15	136549	Homo sapiens transcribed sequences	R34490
AP-hu01	E	16	E16	288663	gap junction protein, beta 1, 32kDa (connexin 32, Charcot-Marie-Tooth neuropathy, X-linked)	N79360
AP-hu01	E	17	E17	136919	chemokine (C-C motif) ligand 28	R38459
AP-hu01	E	18	E18	289268	S100 calcium binding protein A3	N68995
AP-hu01	E	19	E19	137297	LOC151201	R36598
AP-hu01	E	20	E20	289459	pyruvate carboxylase	N76699
AP-hu01	E	21	E21	137995	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	R63185
AP-hu01	E	22	E22	289844	phosphatidylinositol glycan, class A (paroxysmal nocturnal hemoglobinuria)	N77115
AP-hu01	E	23	E23	139125	0	R62517
AP-hu01	E	24	E24	289981	0	N79966
AP-hu01	F	1	F1	415818	LOC253392	W84813
AP-hu01	F	2	F2	756596	dual specificity phosphatase 4	AA444068
AP-hu01	F	3	F3	415851	ras homolog gene family, member E	W86326
AP-hu01	F	4	F4	756876	solute carrier family 35 (UDP-galactose transporter), member A2	AA428423
AP-hu01	F	5	F5	416803	Homo sapiens transcribed sequences	W86776
AP-hu01	F	6	F6	756965	regulator of G-protein signalling 14	AA428897
AP-hu01	F	7	F7	417258	fibrinogen-like 2	W87943
AP-hu01	F	8	F8	757718	nuclear transcription factor Y, gamma	AA442448
AP-hu01	F	9	F9	417361	cell growth regulatory with ring finger domain	W89211
AP-hu01	F	10	F10	757875	cyclin-dependent kinase 5, regulatory subunit 1 (p35)	AA442854
AP-hu01	F	11	F11	417435	hypothetical protein BC013995	W89047
AP-hu01	F	12	F12	757968	transcriptional activator of the c-fos promoter	AA442633
AP-hu01	F	13	F13	417801	mitochondrial ribosomal protein L27	W88848
AP-hu01	F	14	F14	758534	density-regulated protein	AA403237
AP-hu01	F	15	F15	418033	carboxypeptidase A1 (pancreatic)	W90535
AP-hu01	F	16	F16	759207	excision repair cross-complementing rodent repair deficiency, complementation group 4	AA496054
AP-hu01	F	17	F17	418040	TGFB-induced factor (TALE family homeobox)	W90532
AP-hu01	F	18	F18	759881	fibrillin 1 (Marfan syndrome)	AA423948
AP-hu01	F	19	F19	427750	spectrin, alpha, erythrocytic 1 (elliptocytosis 2)	AA001897
AP-hu01	F	20	F20	767174	Purkinje cell protein 4	AA424561
AP-hu01	F	21	F21	428412	granzyme K (serine protease, granzyme 3; tryptase II)	AA005382
AP-hu01	F	22	F22	767295	neurofibromin 2 (bilateral acoustic neuroma)	AA418463
AP-hu01	F	23	F23	470958	hypothetical protein FLJ10420	AA033649
AP-hu01	F	24	F24	767302	KIAA0251 protein	AA418455
AP-hu01	G	1	G1	140197	neuritin 1	R67704
AP-hu01	G	2	G2	290340	Homo sapiens transcribed sequences	N64489
AP-hu01	G	3	G3	140536	neuropilin 2	R67834
AP-hu01	G	4	G4	290517	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	N79133
AP-hu01	G	5	G5	141815	jagged 1 (Alagille syndrome)	R70685
AP-hu01	G	6	G6	290549	Gem-interacting protein	AF086104
AP-hu01	G	7	G7	142383	hypothetical protein FLJ35036	R69925
AP-hu01	G	8	G8	291426	acetyl-Coenzyme A synthetase 2 (AMP forming)-like	W03611
AP-hu01	G	9	G9	142944	chromosome 6 open reading frame 37	R71124
AP-hu01	G	10	G10	292212	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	N80617
AP-hu01	G	11	G11	143887	protease, serine, 23	R76394
AP-hu01	G	12	G12	292699	ankyrin 1, erythrocytic	N80475
AP-hu01	G	13	G13	144797	a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 1	R76553
AP-hu01	G	14	G14	293017	chromodomain helicase DNA binding protein 2	N90649
AP-hu01	G	15	G15	145284	nuclear receptor coactivator 2	R77864
AP-hu01	G	16	G16	293914	sperm associated antigen 9	N66027
AP-hu01	G	17	G17	145696	killer cell lectin-like receptor subfamily D, member 1	R78286

AP-hu01	G	18	G18	298149	fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)	W00709
AP-hu01	G	19	G19	151842	hypothetical protein FLJ12684	H03070
AP-hu01	G	20	G20	298601	likely ortholog of mouse phosphorylated adaptor for RNA export	W04927
AP-hu01	G	21	G21	154742	tumor necrosis factor (ligand) superfamily, member 12	R55379
AP-hu01	G	22	G22	300071	thyroid transcription factor 1	N91550
AP-hu01	G	23	G23	155766	Homo sapiens transcribed sequences	R72149
AP-hu01	G	24	G24	300411	chemokine (C-C motif) ligand 19	W07401
AP-hu01	H	1	H1	471580	BCL2-associated athanogene 3	AA034946
AP-hu01	H	2	H2	767746	similar to ICEBERG caspase-1 inhibitor	AA418193
AP-hu01	H	3	H3	471639	chromosome 2 open reading frame 3	AA035031
AP-hu01	H	4	H4	767808	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	AA418709
AP-hu01	H	5	H5	485294	programmed death ligand 2	AA039608
AP-hu01	H	6	H6	768064	cytochrome P450, family 1, subfamily A, polypeptide 1	AA426553
AP-hu01	H	7	H7	485674	isocitrate dehydrogenase 1 (NADP+), soluble	AA039874
AP-hu01	H	8	H8	768067	ribonuclease, RNase A family, 4	AA426582
AP-hu01	H	9	H9	485778	Sapiens mRNA; cDNA DKFZp564O1016 (from clone DKFZp564O1016)	AA040319
AP-hu01	H	10	H10	768617	diphtheria toxin receptor (heparin-binding epidermal growth factor-like growth factor)	AA430318
AP-hu01	H	11	H11	485857	delta sleep inducing peptide, immunoreactor	AA040063
AP-hu01	H	12	H12	770879	8D6 antigen	AA434501
AP-hu01	H	13	H13	485886	carnitine acetyltransferase	AA040121
AP-hu01	H	14	H14	770989	reticulon 4 receptor	AA428418
AP-hu01	H	15	H15	486233	RAB7, member RAS oncogene family	AA043680
AP-hu01	H	16	H16	771258	CD8 antigen, alpha polypeptide (p32)	AA443649
AP-hu01	H	17	H17	486660	myeloid cell leukemia sequence 1 (BCL2-related)	AA044316
AP-hu01	H	18	H18	771312	testis zinc finger protein	AA476210
AP-hu01	H	19	H19	486696	chemokine (C-X-C motif) receptor 6	AA044449
AP-hu01	H	20	H20	773200	lectin, mannose-binding, 1	AA428522
AP-hu01	H	21	H21	486727	DnaJ (Hsp40) homolog, subfamily A, member 4	AA044432
AP-hu01	H	22	H22	773399	dystrophin (muscular dystrophy, Duchenne and Becker types)	AA425649
AP-hu01	H	23	H23	487396	Sapiens mRNA; cDNA DKFZp564F053 (from clone DKFZp564F053)	AA046660
AP-hu01	H	24	H24	773494	cathepsin K (pycnodysostosis)	AA428008
AP-hu01	I	1	I1	160263	hypothetical protein FLJ22635	H21954
AP-hu01	I	2	I2	301740	hypothetical protein LOC139202	W17094
AP-hu01	I	3	I3	160946	c-fos induced growth factor (vascular endothelial growth factor D)	H24828
AP-hu01	I	4	I4	301851	platelet factor 4 (chemokine (C-X-C motif) ligand 4)	W17158
AP-hu01	I	5	I5	161323	autism susceptibility candidate 2	H25394
AP-hu01	I	6	I6	302052	adaptor-related protein complex 3, mu 2 subunit	N89763
AP-hu01	I	7	I7	162077	pleckstrin homology-like domain, family A, member 1	H26271
AP-hu01	I	8	I8	302127	lipocalin 2 (oncogene 24p3)	W38398
AP-hu01	I	9	I9	163838	protein kinase C, gamma	H14047
AP-hu01	I	10	I10	302188	nebulette	W16723
AP-hu01	I	11	I11	174835	microtubule-associated protein 1B	AF085863
AP-hu01	I	12	I12	305284	hydroxysteroid (17-beta) dehydrogenase 12	N95053
AP-hu01	I	13	I13	176198	ataxin 2-binding protein 1	H41778
AP-hu01	I	14	I14	306813	v-ros UR2 sarcoma virus oncogene homolog 1 (avian)	W24201
AP-hu01	I	15	I15	177341	ankyrin repeat and SOCS box-containing 3	H40739
AP-hu01	I	16	I16	307506	Meis1, myeloid ecotropic viral integration site 1 homolog (mouse)	W21073
AP-hu01	I	17	I17	178257	0	H46830
AP-hu01	I	18	I18	308788	phosphorylase kinase, gamma 1 (muscle)	N93274
AP-hu01	I	19	I19	191708	protein tyrosine phosphatase, non-receptor type 3	H40246
AP-hu01	I	20	I20	310070	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 16	N98554
AP-hu01	I	21	I21	196005	aquaporin 3	R91904
AP-hu01	I	22	I22	322939	forkhead box D1	W45110
AP-hu01	I	23	I23	196387	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	R91503
AP-hu01	I	24	I24	323078	heme oxygenase (decycling) 1	W42558
AP-hu01	J	1	J1	487836	nuclear phosphoprotein similar to S. cerevisiae PWP1	AA045396
AP-hu01	J	2	J2	773568	POU domain, class 4, transcription factor 1	AA428894
AP-hu01	J	3	J3	487965	ubiquitin-conjugating enzyme E2A (RAD6 homolog)	AA045751
AP-hu01	J	4	J4	781738	GATA binding protein 4	AA431631
AP-hu01	J	5	J5	488033	DnaJ (Hsp40) homolog, subfamily B, member 9	AA045793
AP-hu01	J	6	J6	782288	glutathione S-transferase A2	AA432244
AP-hu01	J	7	J7	488629	hypothetical protein FLJ10154	AA045908
AP-hu01	J	8	J8	782758	CD209 antigen-like	AA448002
AP-hu01	J	9	J9	488842	osteoblast specific factor 2 (fasciclin I-like)	AA046245
AP-hu01	J	10	J10	784283	0	AA447502
AP-hu01	J	11	J11	489042	tumor necrosis factor receptor superfamily, member 8	AA047162
AP-hu01	J	12	J12	784585	butyrylcholinesterase	AA443299
AP-hu01	J	13	J13	489212	pyruvate dehydrogenase (lipoamide) alpha 1	AA045659
AP-hu01	J	14	J14	784957	REV3-like, catalytic subunit of DNA polymerase zeta (yeast)	AA448067
AP-hu01	J	15	J15	490004	cytokine receptor-like factor 1	AA121532
AP-hu01	J	16	J16	786125	TAF12 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 20kDa	AA447903
AP-hu01	J	17	J17	490023	wingless-type MMTV integration site family, member 5B	AA114966
AP-hu01	J	18	J18	789049	Sapiens full length insert cDNA clone YW26E10	AA453077
AP-hu01	J	19	J19	490111	O-6-methylguanine-DNA methyltransferase	AA137218
AP-hu01	J	20	J20	795562	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	AA459795
AP-hu01	J	21	J21	490642	cullin 3	AA115733
AP-hu01	J	22	J22	795871	muscleblind-like 3 (Drosophila)	AA460747
AP-hu01	J	23	J23	501994	interleukin 1 receptor-like 1	AA128153
AP-hu01	J	24	J24	796199	hypothetical protein FLJ20069	AA461119
AP-hu01	K	1	K1	197105	KIAA0265 protein	R93904
AP-hu01	K	2	K2	323151	nuclear receptor subfamily 4, group A, member 3	W42610
AP-hu01	K	3	K3	201268	ephrin-A5	R99405
AP-hu01	K	4	K4	323181	fibroblast activation protein, alpha	W45237
AP-hu01	K	5	K5	202112	nuclear factor (erythroid-derived 2), 45kDa	H48815
AP-hu01	K	6	K6	323238	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	W42812
AP-hu01	K	7	K7	202603	BCL2-associated athanogene 4	H53289

AP-hu01	K	8	K8	324054	nerve growth factor, beta polypeptide	W46522
AP-hu01	K	9	K9	203514	hypothetical protein FLJ00225	H55966
AP-hu01	K	10	K10	325365	HIV-1 rev binding protein 2	W52273
AP-hu01	K	11	K11	205898	p8 protein (candidate of metastasis 1)	H58465
AP-hu01	K	12	K12	327060	glutathione S-transferase M1	W25634
AP-hu01	K	13	K13	206779	Homo sapiens transcribed sequences	R98072
AP-hu01	K	14	K14	340588	uridine monophosphate synthetase (orotate phosphoribosyl transferase and orotidine-5'-decarboxylase)	W56655
AP-hu01	K	15	K15	206917	adenylate kinase 3	R98700
AP-hu01	K	16	K16	340644	integrin, beta 8	W56754
AP-hu01	K	17	K17	207416	erythrocyte membrane protein band 4.2	H58859
AP-hu01	K	18	K18	341295	inducible T-cell co-stimulator	W58028
AP-hu01	K	19	K19	207665	Homo sapiens transcribed sequence with weak similarity to protein ref:NP_060434.1 (H.sapiens) hypothetical protein FLJ20837 [Homo sapiens]	H62267
AP-hu01	K	20	K20	341763	caspase 5, apoptosis-related cysteine protease	W60764
AP-hu01	K	21	K21	207838	hypothetical protein LOC129607	H60298
AP-hu01	K	22	K22	342254	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	W61199
AP-hu01	K	23	K23	209731	pim-1 oncogene	H52219
AP-hu01	K	24	K24	342557	complement C1r-like proteinase	W68477
AP-hu01	L	1	L1	502055	arylsulfatase B	AA128826
AP-hu01	L	2	L2	796702	FKBP-associated protein	AA443131
AP-hu01	L	3	L3	503281	interleukin 19	AA151736
AP-hu01	L	4	L4	809464	fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, Jackson-Weiss syndrome)	AA456160
AP-hu01	L	5	L5	503313	interferon-induced protein with tetratricopeptide repeats 2	AA130252
AP-hu01	L	6	L6	810391	hyaluronoglucosaminidase 1	AA464196
AP-hu01	L	7	L7	503343	collagen, type IV, alpha 1	AA130288
AP-hu01	L	8	L8	811942	general transcription factor IIH, polypeptide 1, 62kDa	AA455004
AP-hu01	L	9	L9	504535	SEC24 related gene family, member C (S. cerevisiae)	AA150049
AP-hu01	L	10	L10	814287	X-ray repair complementing defective repair in Chinese hamster cells 3	AA459237
AP-hu01	L	11	L11	505053	leptin receptor gene-related protein	AA149805
AP-hu01	L	12	L12	814578	matrix metalloproteinase 19	AA480939
AP-hu01	L	13	L13	665121	aldehyde dehydrogenase 3 family, member B1	AA194963
AP-hu01	L	14	L14	815758	epidermal growth factor receptor pathway substrate 15	AA485200
AP-hu01	L	15	L15	665424	histidyl-tRNA synthetase-like	AA195056
AP-hu01	L	16	L16	824863	0	AA488959
AP-hu01	L	17	L17	666445	BRF1 homolog, subunit of RNA polymerase III transcription initiation factor IIIB (S. cerevisiae)	AA232937
AP-hu01	L	18	L18	825287	tumor necrosis factor (ligand) superfamily, member 11	AA504450
AP-hu01	L	19	L19	667290	tyrosinase-related protein 1	AA227492
AP-hu01	L	20	L20	825352	histone deacetylase 9	AA504568
AP-hu01	L	21	L21	667407	sperm associated antigen 9	AA227958
AP-hu01	L	22	L22	878499	hemoglobin, beta	AA775838
AP-hu01	L	23	L23	667440	B aggressive lymphoma gene	AA228088
AP-hu01	L	24	L24	1030527	general transcription factor IIB	AA555198
AP-hu01	M	1	M1	210587	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D	H65936
AP-hu01	M	2	M2	342614	cysteine and glycine-rich protein 3 (cardiac LIM protein)	W68761
AP-hu01	M	3	M3	210658	0	H66101
AP-hu01	M	4	M4	342633	caspase recruitment domain family, member 10	W68684
AP-hu01	M	5	M5	212088	ATP-binding cassette, sub-family G (WHITE), member 1	H68960
AP-hu01	M	6	M6	343111	pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)	W67392
AP-hu01	M	7	M7	213414	serine (or cysteine) proteinase inhibitor, clade D (heparin cofactor), member 1	H72143
AP-hu01	M	8	M8	343311	nerve growth factor receptor (TNFR superfamily, member 16)	W67446
AP-hu01	M	9	M9	214144	Rho-associated, coiled-coil containing protein kinase 2	H70847
AP-hu01	M	10	M10	343698	calcium/calmodulin-dependent protein kinase kinase 1, alpha	W69176
AP-hu01	M	11	M11	214427	hypothetical protein MGC4309	H77955
AP-hu01	M	12	M12	344107	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	W73784
AP-hu01	M	13	M13	214469	formin binding protein 3	H73833
AP-hu01	M	14	M14	344384	cell death-inducing DFFA-like effector a	W73447
AP-hu01	M	15	M15	219766	echinoderm microtubule associated protein like 1	H81618
AP-hu01	M	16	M16	344443	thioredoxin-like 2	W73191
AP-hu01	M	17	M17	220826	0	H95623
AP-hu01	M	18	M18	344889	0	W76009
AP-hu01	M	19	M19	221191	0	H91887
AP-hu01	M	20	M20	344953	lymphoid enhancer-binding factor 1	AF086339
AP-hu01	M	21	M21	221527	Homo sapiens transcribed sequence with moderate similarity to protein ref:NP_115694.1 (H.sapiens) hypothetical protein MGC12945 [Homo sapiens]	H92126
AP-hu01	M	22	M22	345034	chemokine (C-X-C motif) ligand 14	AF086340
AP-hu01	M	23	M23	222454	leukocyte membrane antigen	H84077
AP-hu01	M	24	M24	345214	apoptosis antagonizing transcription factor	W72325
AP-hu01	N	1	N1	667596	myogenic factor 6 (herculin)	AA228129
AP-hu01	N	2	N2	1046762	KIAA0562 gene product	AA644350
AP-hu01	N	3	N3	668487	TAF9 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 32kDa	AA243818
AP-hu01	N	4	N4	1075967	BCL2 binding component 3	AI821703
AP-hu01	N	5	N5	669163	translocation protein 1	AA234058
AP-hu01	N	6	N6	1146395	SH2 domain protein 2A	AA613070
AP-hu01	N	7	N7	669194	zinc finger protein 451	AA236470
AP-hu01	N	8	N8	1147276	homeo box A9	AA627519
AP-hu01	N	9	N9	669496	KIAA0763 gene product	AA234778
AP-hu01	N	10	N10	1184337	Homo sapiens transcribed sequence with weak similarity to protein ref:NP_060312.1 (H.sapiens) hypothetical protein FLJ20489 [Homo sapiens]	AA648413
AP-hu01	N	11	N11	681948	fragile histidine triad gene	AA256123
AP-hu01	N	12	N12	1185189	fragile X mental retardation 1	AA648969
AP-hu01	N	13	N13	683915	centaurin, beta 2	AA237088
AP-hu01	N	14	N14	1185311	glia maturation factor, beta	AA651767
AP-hu01	N	15	N15	685336	Sapiens cDNA FLJ36913 fis, clone BRACE2003848, highly similar to Human	AA261775

					SEC14L mRNA.	
AP-hu01	N	16	N16	1251853	KIAA0753 gene product	AA731225
AP-hu01	N	17	N17	686199	KIAA0053 gene product	AA262297
AP-hu01	N	18	N18	1270940	Homo sapiens transcribed sequence with moderate similarity to protein pir:TVHUVV (H.sapiens) TVHUVV transforming protein vav - human (fragments)	AA748610
AP-hu01	N	19	N19	686216	apolipoprotein L, 3	AA262166
AP-hu01	N	20	N20	1307938	T-cell acute lymphocytic leukemia 2	AA767775
AP-hu01	N	21	N21	686463	KIAA0582 protein	AA255947
AP-hu01	N	22	N22	1318267	S100 calcium binding protein A5	AA810909
AP-hu01	N	23	N23	687208	DNA cross-link repair 1C (PSO2 homolog, S. cerevisiae)	AA259032
AP-hu01	N	24	N24	1322986	thiopurine S-methyltransferase	AA746081
AP-hu01	O	1	O1	230126	vitronectin (serum spreading factor, somatomedin B, complement S-protein)	H78801
AP-hu01	O	2	O2	345858	cisplatin resistance associated	W77812
AP-hu01	O	3	O3	233514	Homo sapiens transcribed sequences	H77320
AP-hu01	O	4	O4	346016	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide I	W72635
AP-hu01	O	5	O5	233939	NPC-related protein NAG73	AF085991
AP-hu01	O	6	O6	346146	branched chain keto acid dehydrogenase E1, alpha polypeptide (maple syrup urine disease)	W78018
AP-hu01	O	7	O7	242643	macrophage expressed gene 1	H94982
AP-hu01	O	8	O8	346604	advanced glycosylation end product-specific receptor	W74536
AP-hu01	O	9	O9	242770	metal-regulatory transcription factor 1	H94113
AP-hu01	O	10	O10	346628	S100 calcium binding protein A8 (calgranulin A)	W74581
AP-hu01	O	11	O11	243341	CCCTC-binding factor (zinc finger protein)	N48105
AP-hu01	O	12	O12	346653	chromosome 5 open reading frame 13	W74451
AP-hu01	O	13	O13	243526	KIAA0241 protein	N49464
AP-hu01	O	14	O14	346661	enolase 3, (beta, muscle)	W94187
AP-hu01	O	15	O15	243554	Homo sapiens transcribed sequence with strong similarity to protein pdb:1FDH (H.sapiens) G Chain G, Hemoglobin (Deoxy, Human Fetal FII)	N49794
AP-hu01	O	16	O16	346832	metallothionein 1G	W78010
AP-hu01	O	17	O17	243741	gamma-aminobutyric acid (GABA) B receptor, 1	N49629
AP-hu01	O	18	O18	347092	COX17 homolog, cytochrome c oxidase assembly protein (yeast)	W79621
AP-hu01	O	19	O19	244970	PHD finger protein 3	N76260
AP-hu01	O	20	O20	347456	laminin, alpha 4	W81662
AP-hu01	O	21	O21	245015	NICE-4 protein	N72374
AP-hu01	O	22	O22	347736	solute carrier family 35, member E2	W81638
AP-hu01	O	23	O23	245569	solute carrier family 22 (extraneuronal monoamine transporter), member 3	N77239
AP-hu01	O	24	O24	347761	src family associated phosphoprotein 2	W81605
AP-hu01	P	1	P1	687234	casein kinase 1, epsilon	AA257959
AP-hu01	P	2	P2	1327016	SP110 nuclear body protein	AA724854
AP-hu01	P	3	P3	687592	oncostatin M receptor	AA234544
AP-hu01	P	4	P4	1335756	Homo sapiens transcribed sequences	AA828293
AP-hu01	P	5	P5	700393	solute carrier family 35 (CMP-sialic acid transporter), member A1	AA290699
AP-hu01	P	6	P6	1336836	RNA binding motif protein 15	AA809402
AP-hu01	P	7	P7	700668	glycine receptor, alpha 3	AA283885
AP-hu01	P	8	P8	1339205	B-cell CLL/lymphoma 11A (zinc finger protein)	AA814477
AP-hu01	P	9	P9	701127	basic leucine zipper nuclear factor 1 (JEM-1)	AA287458
AP-hu01	P	10	P10	1339831	regulatory factor X, 5 (influences HLA class II expression)	AA815144
AP-hu01	P	11	P11	701272	T-cell leukemia/lymphoma 6	AA287767
AP-hu01	P	12	P12	1339933	apoptosis-related protein PNAS-1	AA815207
AP-hu01	P	13	P13	704600	tuberous sclerosis 1	AA279895
AP-hu01	P	14	P14	1352486	Homo sapiens transcribed sequences	AA829994
AP-hu01	P	15	P15	705056	butyrophilin, subfamily 1, member A1	AA279987
AP-hu01	P	16	P16	1370282	SMC5 structural maintenance of chromosomes 5-like 1 (yeast)	AA847845
AP-hu01	P	17	P17	712739	similar to S. cerevisiae SSM4	AA280560
AP-hu01	P	18	P18	1370624	hypothetical protein LOC339091	AA835686
AP-hu01	P	19	P19	713862	ubiquitin-protein isopeptide ligase (E3)	AA284599
AP-hu01	P	20	P20	1371906	SKI-like	AA827643
AP-hu01	P	21	P21	725269	secreted frizzled-related protein 4	AA291847
AP-hu01	P	22	P22	1371945	Glucocorticoid receptor alpha mRNA, variant 3' UTR	AA827786
AP-hu01	P	23	P23	726563	transcription elongation factor A (SII), 2	AA394221
AP-hu01	P	24	P24	1388373	0	AA844124
AP-hu02	A	1	A1	1401793	HUS1 checkpoint homolog (S. pombe)	AA854956
AP-hu02	A	2	A2	2150232	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	AI457252
AP-hu02	A	3	A3	1406233	transmembrane 4 superfamily member 1	AA889018
AP-hu02	A	4	A4	2150502	chemokine (C-C motif) ligand 22	AI457797
AP-hu02	A	5	A5	1411935	Sapiens, Similar to phospholipase A2, group IB (pancreas), clone MGC:12509 IMAGE:3950450, mRNA, complete cds	AA921407
AP-hu02	A	6	A6	2157583	apoptosis regulator BCL-G	AI478889
AP-hu02	A	7	A7	1466849	renal tumor antigen	AA885446
AP-hu02	A	8	A8	2162740	Fas apoptotic inhibitory molecule 2	AI468058
AP-hu02	A	9	A9	1467763	0	AA883106
AP-hu02	A	10	A10	2167853	phosphoinositide-3-kinase adaptor protein	AI582746
AP-hu02	A	11	A11	1469157	host cell factor C1 (VP16-accessory protein)	AA865726
AP-hu02	A	12	A12	2237374	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1	AI623172
AP-hu02	A	13	A13	1472204	complement component 1, s subcomponent	AA873271
AP-hu02	A	14	A14	2242064	Homo sapiens transcribed sequence with strong similarity to protein sp:P25391 (H.sapiens) LMA1_HUMAN Laminin alpha-1 chain precursor (Laminin A chain)	AI633741
AP-hu02	A	15	A15	1507192	AT-hook transcription factor AKNA	AA904466
AP-hu02	A	16	A16	2244541	G protein-coupled receptor 86	AI656746
AP-hu02	A	17	A17	1520827	HIV TAT specific factor 1	AA910247
AP-hu02	A	18	A18	2248373	I-kappa-B-interacting Ras-like protein 1	AI659150
AP-hu02	A	19	A19	1521793	tumor necrosis factor (ligand) superfamily, member 13b	AA906714
AP-hu02	A	20	A20	2271560	insulinoma-associated 1	AI680827
AP-hu02	A	21	A21	1522492	postmeiotic segregation increased 2-like 6	AA908680
AP-hu02	A	22	A22	2273055	Homo sapiens transcribed sequence with weak similarity to protein sp:P11369 (M.musculus) POL2_MOUSE Retrovirus-related POL polyprotein [Contains: Reverse transcriptase ; Endonuclease]	AI681945
AP-hu02	A	23	A23	1523511	core-binding factor, runt domain, alpha subunit 2; translocated to, 1; cyclin D-related	AA904353

AP-hu02	A	24	A24	2285313	guanine nucleotide binding protein (G protein), alpha 14	AI629022
AP-hu02	B	1	B1	470185	mitogen-activated protein kinase kinase kinase 7 interacting protein 1	AA028980
AP-hu02	B	2	B2	1144761	low density lipoprotein receptor (familial hypercholesterolemia)	AA621897
AP-hu02	B	3	B3	471252	retinoic acid receptor, gamma	AA034461
AP-hu02	B	4	B4	1183847	syndecan binding protein (syntenin)	AA687473
AP-hu02	B	5	B5	486128	fibrillin 1 (Marfan syndrome)	AA043236
AP-hu02	B	6	B6	1258499	interleukin 23, alpha subunit p19	AA729815
AP-hu02	B	7	B7	487402	polymerase (DNA directed), mu	AA046663
AP-hu02	B	8	B8	1272059	granzyme K (serine protease, granzyme 3; tryptase II)	AA743283
AP-hu02	B	9	B9	488909	histone deacetylase 5	AA057030
AP-hu02	B	10	B10	1285738	Homo sapiens transcribed sequence with weak similarity to protein ref:NP_060312.1 (H.sapiens) hypothetical protein FLJ20489 [Homo sapiens]	AA743682
AP-hu02	B	11	B11	489860	serologically defined colon cancer antigen 8	AA102109
AP-hu02	B	12	B12	1286225	breast cancer 2, early onset	AA741449
AP-hu02	B	13	B13	489870	Sapiens, clone IMAGE:5535936, mRNA	AA114941
AP-hu02	B	14	B14	1288031	SMC5 structural maintenance of chromosomes 5-like 1 (yeast)	AA809572
AP-hu02	B	15	B15	491677	integrin, beta-like 1 (with EGF-like repeat domains)	AA150407
AP-hu02	B	16	B16	1301819	oxidase (cytochrome c) assembly 1-like	AA767234
AP-hu02	B	17	B17	491727	Homo sapiens transcribed sequence with weak similarity to protein ref:NP_060265.1 (H.sapiens) hypothetical protein FLJ20378 [Homo sapiens]	AA156691
AP-hu02	B	18	B18	1305734	Glucocorticoid receptor alpha mRNA, variant 3' UTR	AA736674
AP-hu02	B	19	B19	501883	fibromodulin	AA129908
AP-hu02	B	20	B20	1316966	Homo sapiens transcribed sequence with strong similarity to protein pir:CRHU5 (H.sapiens) CRHU5 carbonate dehydratase (EC 4.2.1.1) V precursor - human	AA767841
AP-hu02	B	21	B21	502257	cathepsin K (pycnodysostosis)	AA156664
AP-hu02	B	22	B22	1321372	caspase recruitment domain family, member 10	AA759191
AP-hu02	B	23	B23	502463	zinc finger protein 24 (KOX 17)	AA134771
AP-hu02	B	24	B24	1324864	0	AA740506
AP-hu02	C	1	C1	1524563	S-antigen; retina and pineal gland (arrestin)	AA904507
AP-hu02	C	2	C2	2285718	cytochrome P450, family 3, subfamily A, polypeptide 4	AI758923
AP-hu02	C	3	C3	1540470	ubiquitin carrier protein	AA928110
AP-hu02	C	4	C4	2297925	EphA4	AI913604
AP-hu02	C	5	C5	1542260	pyruvate dehydrogenase (lipoamide) alpha 2	AA927400
AP-hu02	C	6	C6	2303920	death associated transcription factor 1	AI650841
AP-hu02	C	7	C7	1543844	19A24 protein	AA921765
AP-hu02	C	8	C8	2304135	aminoadipate-semialdehyde synthase	AI651132
AP-hu02	C	9	C9	1551765	interleukin 1, alpha	AA922689
AP-hu02	C	10	C10	2315207	chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	AI668847
AP-hu02	C	11	C11	1557028	regulator of G-protein signalling 9	AA915931
AP-hu02	C	12	C12	2318450	C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 12	AI708851
AP-hu02	C	13	C13	1558715	platelet-derived growth factor alpha polypeptide	AA977022
AP-hu02	C	14	C14	2325602	cadherin 17, LI cadherin (liver-intestine)	AI688206
AP-hu02	C	15	C15	1561976	homeo box C11	AA953456
AP-hu02	C	16	C16	2325850	solute carrier family 6 (neurotransmitter transporter, taurine), member 6	AI688443
AP-hu02	C	17	C17	1566020	SCAN domain containing 1	AI073530
AP-hu02	C	18	C18	2330876	complement component 1, q subcomponent, receptor 1	AI693483
AP-hu02	C	19	C19	1570420	chemokine (C-C motif) ligand 1	AA931884
AP-hu02	C	20	C20	2341256	aryl hydrocarbon receptor nuclear translocator	AI697689
AP-hu02	C	21	C21	1579639	interferon, gamma	AA969504
AP-hu02	C	22	C22	2348576	killer cell lectin-like receptor subfamily C, member 2	AI798774
AP-hu02	C	23	C23	1581406	CD80 antigen (CD28 antigen ligand 1, B7-1 antigen)	AA983817
AP-hu02	C	24	C24	2349996	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	AI806456
AP-hu02	D	1	D1	503206	immunoglobulin mu binding protein 2	AA148925
AP-hu02	D	2	D2	1325272	cytochrome c oxidase subunit VIIc	AA741399
AP-hu02	D	3	D3	503280	interleukin 19	AA151733
AP-hu02	D	4	D4	1326381	wingless-type MMTV integration site family, member 10A	AA759288
AP-hu02	D	5	D5	503746	chromosome 6 open reading frame 4	AA131463
AP-hu02	D	6	D6	1326650	c-fos induced growth factor (vascular endothelial growth factor D)	AA813617
AP-hu02	D	7	D7	504826	transcription factor A, mitochondrial	AA150777
AP-hu02	D	8	D8	1335806	CUG triplet repeat, RNA binding protein 1	AA828368
AP-hu02	D	9	D9	505258	cytochrome P450, family 11, subfamily A, polypeptide 1	AA142879
AP-hu02	D	10	D10	1339537	butyrophilin, subfamily 1, member A1	AA805147
AP-hu02	D	11	D11	667009	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	AA253109
AP-hu02	D	12	D12	1340493	formin binding protein 3	AA938206
AP-hu02	D	13	D13	667614	histone deacetylase 6	AA228073
AP-hu02	D	14	D14	1350983	glia maturation factor, beta	AA806487
AP-hu02	D	15	D15	668465	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)-like 1	AA243734
AP-hu02	D	16	D16	1352169	0	AA808204
AP-hu02	D	17	D17	668977	suppressor of cytokine signaling 4	AA253351
AP-hu02	D	18	D18	1355977	SET domain, bifurcated 2	AA831594
AP-hu02	D	19	D19	669261	Sapiens, clone IMAGE:4214177, mRNA	AA236520
AP-hu02	D	20	D20	1357778	ubiquitin-protein isopeptide ligase (E3)	AA832340
AP-hu02	D	21	D21	669310	mitogen-activated protein kinase-activated protein kinase 5	AA236814
AP-hu02	D	22	D22	1371215	discs, large (Drosophila) homolog 1	AA847183
AP-hu02	D	23	D23	683358	Homo sapiens transcribed sequences	AA215307
AP-hu02	D	24	D24	1372313	receptor-interacting serine-threonine kinase 2	AA826052
AP-hu02	E	1	E1	1584882	ankyrin repeat and SOCS box-containing 4	AI791680
AP-hu02	E	2	E2	2354198	macrophage scavenger receptor 1	AI734941
AP-hu02	E	3	E3	1586516	phosphoinositide-3-kinase adaptor protein	AA973777
AP-hu02	E	4	E4	2357389	CD1B antigen, b polypeptide	AI827808
AP-hu02	E	5	E5	1589786	fibroblast growth factor 9 (glia-activating factor)	AA946776
AP-hu02	E	6	E6	2359664	Cas-Br-M (murine) ecotropic retroviral transforming sequence	AI808688
AP-hu02	E	7	E7	1602947	chemokine (C-C motif) ligand 8	AA989086
AP-hu02	E	8	E8	2369305	lymphocyte alpha-kinase	AI760166
AP-hu02	E	9	E9	1605126	leukotriene A4 hydrolase	AA987920
AP-hu02	E	10	E10	2370171	A kinase (PRKA) anchor protein (yotiao) 9	AI761727

AP-hu02	E	11	E11	1621044	ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome	AI004861
AP-hu02	E	12	E12	2379167	vanin 1	AI762340
AP-hu02	E	13	E13	1625719	protein tyrosine phosphatase, receptor type, D	AI005089
AP-hu02	E	14	E14	2382863	cytochrome P450, family 3, subfamily A, polypeptide 5	AI759985
AP-hu02	E	15	E15	1634832	class I cytokine receptor	AI088984
AP-hu02	E	16	E16	2383199	potassium inwardly-rectifying channel, subfamily J, member 1	AI795925
AP-hu02	E	17	E17	1638431	postmeiotic segregation increased 2-like 5	AI015925
AP-hu02	E	18	E18	2384363	0	AI830536
AP-hu02	E	19	E19	1641007	0	AI040940
AP-hu02	E	20	E20	2386044	Kruppel-like factor 2 (lung)	AI765962
AP-hu02	E	21	E21	1643606	protein transport protein SEC61 alpha subunit isoform 1	AI023590
AP-hu02	E	22	E22	2386478	NP220 nuclear protein	AI808939
AP-hu02	E	23	E23	1645371	cytochrome P450, family 17, subfamily A, polypeptide 1	AI028749
AP-hu02	E	24	E24	2392865	granzyme M (lymphocyte met-ase 1)	AI910778
AP-hu02	F	1	F1	683822	ATP-binding cassette, sub-family F (GCN20), member 1	AA255836
AP-hu02	F	2	F2	1373030	Sapiens cDNA FLJ33408 fis, clone BRACE2010550, highly similar to Human cell growth regulator CGR19 mRNA.	AA825685
AP-hu02	F	3	F3	684760	hypothetical protein MGC4175	AA251625
AP-hu02	F	4	F4	1387641	glutathione S-transferase A3	AA846513
AP-hu02	F	5	F5	684917	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)	AA251909
AP-hu02	F	6	F6	1401913	histone deacetylase 7A	AA854147
AP-hu02	F	7	F7	685470	tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator)	AA262421
AP-hu02	F	8	F8	1405643	KIAA0562 gene product	AA890549
AP-hu02	F	9	F9	685554	macrophage stimulating, pseudogene 9	AA262037
AP-hu02	F	10	F10	1405890	I-kappa-B-interacting Ras-like protein 1	AA890397
AP-hu02	F	11	F11	687301	chemokine (C-X-C motif) ligand 16	AA235226
AP-hu02	F	12	F12	1407260	thymic stromal lymphopoietin	AA889581
AP-hu02	F	13	F13	687934	translocation protein 1	AA236805
AP-hu02	F	14	F14	1412597	carboxypeptidase A1 (pancreatic)	AA835150
AP-hu02	F	15	F15	700704	ankyrin repeat and SOCS box-containing 3	AA285075
AP-hu02	F	16	F16	1435368	PRKC, apoptosis, WT1, regulator	AA834474
AP-hu02	F	17	F17	703508	0	AA278673
AP-hu02	F	18	F18	1435643	death associated protein 3	AA858184
AP-hu02	F	19	F19	703643	tripartite motif-containing 14	AA278637
AP-hu02	F	20	F20	1455450	chemokine (C-X-C motif) ligand 14	AA865643
AP-hu02	F	21	F21	704456	deoxyribonuclease I-like 3	AA279835
AP-hu02	F	22	F22	1467195	cytochrome P450, family 11, subfamily B, polypeptide 1	AA884709
AP-hu02	F	23	F23	712161	histone 1, H2bk	AA280257
AP-hu02	F	24	F24	1492385	phosphoinositide-3-kinase adaptor protein	AA894359
AP-hu02	G	1	G1	1651122	ubiquitin specific protease 15	AI091570
AP-hu02	G	2	G2	2394148	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1	AI763334
AP-hu02	G	3	G3	1662143	sphingomyelin phosphodiesterase 2, neutral membrane (neutral sphingomyelinase)	AI083508
AP-hu02	G	4	G4	2404737	T cell receptor alpha locus	AI816828
AP-hu02	G	5	G5	1662507	KIAA0169 protein	AI083541
AP-hu02	G	6	G6	2406104	interleukin 1 family, member 9	AI828999
AP-hu02	G	7	G7	1663139	interleukin 11	AI128942
AP-hu02	G	8	G8	2406442	UL16 binding protein 1	AI830832
AP-hu02	G	9	G9	1663164	0	AI129113
AP-hu02	G	10	G10	2409058	B cell RAG associated protein	AI819384
AP-hu02	G	11	G11	1664589	collagen, type XI, alpha 1	AI041918
AP-hu02	G	12	G12	2429773	frizzled homolog 2 (Drosophila)	AI858773
AP-hu02	G	13	G13	1672339	caspase recruitment domain family, member 15	AI090427
AP-hu02	G	14	G14	2460033	transcription elongation factor A (SII), 1	AI936239
AP-hu02	G	15	G15	1673209	Fc fragment of IgG, low affinity IIIa, receptor for (CD16)	AI141618
AP-hu02	G	16	G16	2474074	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	AI962842
AP-hu02	G	17	G17	1678175	postsynaptic protein CRIPT	AI086377
AP-hu02	G	18	G18	2478677	T-cell lymphoma invasion and metastasis 1	AW003835
AP-hu02	G	19	G19	1697461	Homo sapiens transcribed sequence with strong similarity to protein pir:B46012 (H.sapiens) B46012 pp52/LSP1 homolog 2 - human (fragment)	AI142438
AP-hu02	G	20	G20	2490395	low density lipoprotein receptor-related protein 6	AI972744
AP-hu02	G	21	G21	1703954	adhesion glycoprotein	AI160261
AP-hu02	G	22	G22	2490644	apoptosis inhibitor 5	AI972925
AP-hu02	G	23	G23	1704556	lipoprotein lipase	AI096515
AP-hu02	G	24	G24	2495108	homeo box B6	AW001035
AP-hu02	H	1	H1	712441	succinate dehydrogenase complex, subunit D, integral membrane protein	AA281913
AP-hu02	H	2	H2	1493445	homeo box C5	AA894914
AP-hu02	H	3	H3	712943	polymerase (DNA directed), epsilon 2 (p59 subunit)	AA282380
AP-hu02	H	4	H4	1504500	T-cell acute lymphocytic leukemia 2	AA904907
AP-hu02	H	5	H5	713619	signal recognition particle receptor ('docking protein')	AA283840
AP-hu02	H	6	H6	1504851	protein phosphatase 2 (formerly 2A), regulatory subunit B", alpha	AA906152
AP-hu02	H	7	H7	714050	uridine monophosphate synthetase (orotate phosphoribosyl transferase and orotidine-5'-decarboxylase)	AA284909
AP-hu02	H	8	H8	1508334	jagged 1 (Alagille syndrome)	AA947272
AP-hu02	H	9	H9	723784	peroxisome proliferative activated receptor, delta	AA292445
AP-hu02	H	10	H10	1527121	vacuolar protein sorting 4B (yeast)	AA916562
AP-hu02	H	11	H11	724368	EphA2	AA411065
AP-hu02	H	12	H12	1534940	Bcl-2-associated transcription factor	AA918867
AP-hu02	H	13	H13	726388	tripartite motif-containing 29	AA293775
AP-hu02	H	14	H14	1543210	S-antigen; retina and pineal gland (arrestin)	AA918179
AP-hu02	H	15	H15	728133	chromosome 12 open reading frame 22	AA435633
AP-hu02	H	16	H16	1543859	signal recognition particle 54kDa	AA921832
AP-hu02	H	17	H17	729886	ATP-binding cassette, sub-family C (CFTR/MRP), member 1	AA412168
AP-hu02	H	18	H18	1555042	par-6 partitioning defective 6 homolog alpha (C.elegans)	AA954518
AP-hu02	H	19	H19	730149	transcription elongation factor A (SII), 2	AA412613
AP-hu02	H	20	H20	1562070	pyruvate dehydrogenase (lipoamide) alpha 1	AA933938
AP-hu02	H	21	H21	730845	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	AA417135
AP-hu02	H	22	H22	1562115	Sapiens mRNA; cDNA DKFZp586B0922 (from clone DKFZp586B0922)	AA936359

AP-hu02	H	23	H23	730902	Mdm4, transformed 3T3 cell double minute 4, p53 binding protein (mouse)	AA416645
AP-hu02	H	24	H24	1562173	RNA binding motif protein 15	AA933971
AP-hu02	I	1	I1	1716220	0	AI160585
AP-hu02	I	2	I2	2536533	zinc finger protein, subfamily 1A, 3 (Aiolos)	AW071111
AP-hu02	I	3	I3	1721224	inositol polyphosphate-5-phosphatase, 40kDa	AI151146
AP-hu02	I	4	I4	2544446	angiopoietin 1	AW044503
AP-hu02	I	5	I5	1731840	myoglobin	AI184297
AP-hu02	I	6	I6	2567597	a disintegrin and metalloproteinase domain 15 (metargidin)	AW072364
AP-hu02	I	7	I7	1742445	chemokine (C-C motif) ligand 26	AI186047
AP-hu02	I	8	I8	2605431	protein phosphatase 1, regulatory (inhibitor) subunit 3A (glycogen and sarcoplasmic reticulum binding subunit, skeletal muscle)	AW118744
AP-hu02	I	9	I9	1751372	homeo box A10	AI187278
AP-hu02	I	10	I10	2662365	B7 homolog 3	AW181901
AP-hu02	I	11	I11	1752018	fms-related tyrosine kinase 3	AI150354
AP-hu02	I	12	I12	2663077	recombination activating gene 1	AW183306
AP-hu02	I	13	I13	1755890	regulatory factor X, 3 (influences HLA class II expression)	AI203838
AP-hu02	I	14	I14	2690595	ret finger protein 2	AW237758
AP-hu02	I	15	I15	1847208	adrenomedullin receptor	AI243295
AP-hu02	I	16	I16	2911855	UL16 binding protein 2	AW510737
AP-hu02	I	17	I17	1849100	sialic acid binding Ig-like lectin 7	AI247571
AP-hu02	I	18	I18	2945121	matrix metalloproteinase 20 (enamelysin)	AW593141
AP-hu02	I	19	I19	1854028	Down syndrome cell adhesion molecule	AI243628
AP-hu02	I	20	I20	2945274	Ksp37 protein	AW593206
AP-hu02	I	21	I21	1857875	alpha-1-microglobulin/bikunin precursor	AI246161
AP-hu02	I	22	I22	2975903	interleukin 21 receptor	AW629191
AP-hu02	I	23	I23	1870748	0	AI245910
AP-hu02	I	24	I24	2717070	transforming growth factor, beta receptor I (activin A receptor type II-like kinase, 53kDa)	AW137880
AP-hu02	J	1	J1	730933	protein phosphatase 2, regulatory subunit B (B56), beta isoform	AA421193
AP-hu02	J	2	J2	1565637	HUS1 checkpoint homolog (S. pombe)	AI167268
AP-hu02	J	3	J3	740286	DR1-associated protein 1 (negative cofactor 2 alpha)	AA476942
AP-hu02	J	4	J4	1585823	alpha-fetoprotein	AA976755
AP-hu02	J	5	J5	741974	0	AA405401
AP-hu02	J	6	J6	1587251	paired-like homeodomain transcription factor 2	AA977103
AP-hu02	J	7	J7	744128	cancer/testis antigen 1	AA634317
AP-hu02	J	8	J8	1587390	interleukin 12 receptor, beta 2	AA977194
AP-hu02	J	9	J9	744130	wingless-type MMTV integration site family, member 3	AA634318
AP-hu02	J	10	J10	1588397	annexin A6	AA975225
AP-hu02	J	11	J11	746189	BCL6 co-repressor	AA417833
AP-hu02	J	12	J12	1609989	ankyrin repeat and SOCS box-containing 1	AI001012
AP-hu02	J	13	J13	753215	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	AA406420
AP-hu02	J	14	J14	1620057	core-binding factor, runt domain, alpha subunit 2; translocated to, 1; cyclin D-related	AA992855
AP-hu02	J	15	J15	753734	TATA box binding protein (TBP)-associated factor, RNA polymerase I, C, 110kDa	AA411726
AP-hu02	J	16	J16	1626874	cysteine and glycine-rich protein 3 (cardiac LIM protein)	AI018334
AP-hu02	J	17	J17	753945	discs, large (Drosophila) homolog 5	AA479265
AP-hu02	J	18	J18	1627091	retinoic acid early transcript 1K pseudogene	AI017460
AP-hu02	J	19	J19	754351	hyaluronoglucosaminidase 1	AA436276
AP-hu02	J	20	J20	1627998	leukemia inhibitory factor receptor	AI014497
AP-hu02	J	21	J21	756502	nudix (nucleoside diphosphate linked moiety X)-type motif 1	AA444020
AP-hu02	J	22	J22	1628761	interleukin 1 family, member 7 (zeta)	AI014548
AP-hu02	J	23	J23	758266	thrombospondin 4	AA437064
AP-hu02	J	24	J24	1631832	Rap1 guanine-nucleotide-exchange factor directly activated by cAMP	AI004302
AP-hu02	K	1	K1	1874560	calmodulin 1 (phosphorylase kinase, delta)	AI281356
AP-hu02	K	2	K2	2719958	interleukin 21 receptor	AW205118
AP-hu02	K	3	K3	1875345	transmembrane 4 superfamily member 3	AI274190
AP-hu02	K	4	K4	2720344	Notch homolog 3 (Drosophila)	AW138903
AP-hu02	K	5	K5	1880926	nuclear RNA export factor 1	AI268362
AP-hu02	K	6	K6	2721127	matrix metalloproteinase 28	AW207107
AP-hu02	K	7	K7	1882591	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	AI279006
AP-hu02	K	8	K8	2724452	protein tyrosine phosphatase, receptor type, H	AW291534
AP-hu02	K	9	K9	1899192	chemokine (C-C motif) receptor-like 2	AI288845
AP-hu02	K	10	K10	3032869	chemokine-like receptor 1	AW771926
AP-hu02	K	11	K11	1912380	glutathione peroxidase 2 (gastrointestinal)	AI304515
AP-hu02	K	12	K12	3035088	ketohexokinase (fructokinase)	AW779891
AP-hu02	K	13	K13	1913168	Homo sapiens transcribed sequence with moderate similarity to protein ref:NP_060265.1 (H.sapiens) hypothetical protein FLJ20378 [Homo sapiens]	AI304610
AP-hu02	K	14	K14	2734295	0	AW449304
AP-hu02	K	15	K15	1917820	cytidine deaminase	AI343863
AP-hu02	K	16	K16	2736931	purinergic receptor P2X-like 1, orphan receptor	AW451665
AP-hu02	K	17	K17	1926309	death associated protein 3	AI346132
AP-hu02	K	18	K18	3053174	postmeiotic segregation increased 2-like 6	AW401404
AP-hu02	K	19	K19	1931676	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	AI333219
AP-hu02	K	20	K20	3054559	branched chain alpha-ketoacid dehydrogenase kinase	AW575117
AP-hu02	K	21	K21	1934125	cytochrome P450, family 11, subfamily B, polypeptide 1	AI334352
AP-hu02	K	22	K22	3059780	polymerase (DNA directed), lambda	AW575985
AP-hu02	K	23	K23	1939054	squalene epoxidase	AI341546
AP-hu02	K	24	K24	3061074	ankyrin repeat and SOCS box-containing 1	AW576120
AP-hu02	L	1	L1	773190	KIAA0763 gene product	AA428474
AP-hu02	L	2	L2	1636813	nuclear phosphoprotein similar to S. cerevisiae PWP1	AI017373
AP-hu02	L	3	L3	773274	frizzled-related protein	AA425521
AP-hu02	L	4	L4	1639806	postmeiotic segregation increased 2-like 5	AI025190
AP-hu02	L	5	L5	781222	TGFB1-induced anti-apoptotic factor 1	AA446223
AP-hu02	L	6	L6	1643585	guanine nucleotide binding protein (G protein), alpha activating activity polypeptide, olfactory type	AI024119
AP-hu02	L	7	L7	781575	B aggressive lymphoma gene	AA431459
AP-hu02	L	8	L8	1644971	SKI-like	AI025801
AP-hu02	L	9	L9	784532	phosphofructokinase, muscle	AA447412

AP-hu02	L	10	L10	1647990	pyridoxal (pyridoxine, vitamin B6) kinase	AI033704
AP-hu02	L	11	L11	784726	C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 12	AA478515
AP-hu02	L	12	L12	1649917	N-acetyltransferase 1 (arylamine N-acetyltransferase)	AI022585
AP-hu02	L	13	L13	785141	serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 2	AA476471
AP-hu02	L	14	L14	1652632	CDC6 cell division cycle 6 homolog (S. cerevisiae)	AI052065
AP-hu02	L	15	L15	785345	optic atrophy 1 (autosomal dominant)	AA476505
AP-hu02	L	16	L16	1653944	chemokine (C-C motif) ligand 8	AI085069
AP-hu02	L	17	L17	788566	Purkinje cell protein 4	AA452966
AP-hu02	L	18	L18	1662047	mitogen-activated protein kinase kinase kinase 7 interacting protein 2	AI082790
AP-hu02	L	19	L19	795275	tumor protein p53 inducible protein 3	AA451810
AP-hu02	L	20	L20	1662532	0	AI083552
AP-hu02	L	21	L21	796085	biglycan	AA460811
AP-hu02	L	22	L22	1662927	Homo sapiens transcribed sequence with strong similarity to protein sp:Q9Y2Y9 (H.sapiens) KLFD_HUMAN Krueppel-like factor 13 (Transcription factor BTEB3) (Basic transcription element binding protein 3) (BTE-binding protein 3) (RANTES factor of late activated T lymphocytes-1) (RFLAT-1) (Transcription factor NSLP1) (Novel Sp1-like zinc finger transcription factor 1) (Transcription factor NSLP1)	AI074707
AP-hu02	L	23	L23	796446	laminin, beta 2 (laminin S)	AA460076
AP-hu02	L	24	L24	1663533	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	AI034287
AP-hu02	M	1	M1	1941548	0	AI206322
AP-hu02	M	2	M2	3068359	transient receptor potential cation channel, subfamily V, member 6	AW452394
AP-hu02	M	3	M3	1942634	programmed cell death 1 ligand 1	AI202996
AP-hu02	M	4	M4	3070861	a disintegrin and metalloproteinase domain 8	BF513621
AP-hu02	M	5	M5	1943890	feline sarcoma oncogene	AI199245
AP-hu02	M	6	M6	3071092	adrenergic, beta, receptor kinase 2	BF513491
AP-hu02	M	7	M7	1944395	thromboxane A synthase 1 (platelet, cytochrome P450, family 5, subfamily A)	AI202694
AP-hu02	M	8	M8	2731798	frizzled homolog 9 (Drosophila)	AW297263
AP-hu02	M	9	M9	1946205	jun D proto-oncogene	AI339181
AP-hu02	M	10	M10	3172977	zinc finger protein 145 (Kruppel-like, expressed in promyelocytic leukemia)	BE222963
AP-hu02	M	11	M11	1948222	PPAR binding protein	AI352382
AP-hu02	M	12	M12	3175743	chemokine (C motif) ligand 2	BE218461
AP-hu02	M	13	M13	2012241	leukocyte membrane antigen	AI356722
AP-hu02	M	14	M14	3197927	fer (fps/fes related) tyrosine kinase (phosphoprotein NCP94)	BE464122
AP-hu02	M	15	M15	2029551	stanniocalcin 1	AI367280
AP-hu02	M	16	M16	3212988	activating transcription factor 5	BE467506
AP-hu02	M	17	M17	2032639	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 1	AI793206
AP-hu02	M	18	M18	3213997	diacylglycerol kinase, gamma 90kDa	BE468027
AP-hu02	M	19	M19	2055309	B aggressive lymphoma gene	AI307751
AP-hu02	M	20	M20	3219673	alkaline phosphatase, intestinal	BE550404
AP-hu02	M	21	M21	2055594	interleukin 18 receptor accessory protein	AI307682
AP-hu02	M	22	M22	3268827	interleukin 1 family, member 5 (delta)	BF435174
AP-hu02	M	23	M23	2061969	interleukin 1 family, member 7 (zeta)	AI343258
AP-hu02	M	24	M24	3323544	tumor necrosis factor receptor superfamily, member 9	BF064069
AP-hu02	N	1	N1	796471	secretory leukocyte protease inhibitor (antileukoproteinase)	AA460434
AP-hu02	N	2	N2	1669661	chromosome 1 open reading frame 16	AI057273
AP-hu02	N	3	N3	809587	Rho-specific guanine nucleotide exchange factor p114	AA456630
AP-hu02	N	4	N4	1670127	HIV TAT specific factor 1	AI074508
AP-hu02	N	5	N5	810137	fructose-1,6-bisphosphatase 1	AA464360
AP-hu02	N	6	N6	1671227	general transcription factor IIIA	AI086037
AP-hu02	N	7	N7	810255	G protein-coupled receptor kinase-interactor 1	AA464724
AP-hu02	N	8	N8	1674708	hypothetical protein MGC4643	AI075771
AP-hu02	N	9	N9	810672	8D6 antigen	AA463987
AP-hu02	N	10	N10	1674873	T-cell acute lymphocytic leukemia 1	AI076543
AP-hu02	N	11	N11	810777	dual specificity phosphatase 4	AA481753
AP-hu02	N	12	N12	1675856	KIAA0217 protein	AI078341
AP-hu02	N	13	N13	811810	TAF5-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65kDa	AA463470
AP-hu02	N	14	N14	1676286	general transcription factor IIB	AI075694
AP-hu02	N	15	N15	812139	aldehyde dehydrogenase 5 family, member A1 (succinate-semialdehyde dehydrogenase)	AA456021
AP-hu02	N	16	N16	1677559	chromosome 14 open reading frame 111	AI078579
AP-hu02	N	17	N17	813991	transcriptional activator of the c-fos promoter	AA455702
AP-hu02	N	18	N18	1678249	0	AI142896
AP-hu02	N	19	N19	814234	succinate-CoA ligase, GDP-forming, beta subunit	AA465233
AP-hu02	N	20	N20	1679256	renal tumor antigen	AI066456
AP-hu02	N	21	N21	814358	0	AA459022
AP-hu02	N	22	N22	1680119	interleukin 11	AI079412
AP-hu02	N	23	N23	814651	checkpoint suppressor 1	AA481094
AP-hu02	N	24	N24	1682318	reticulon 4	AI066675
AP-hu02	O	1	O1	2065463	collagen, type VI, alpha 2	AI375553
AP-hu02	O	2	O2	3072364	zinc finger protein 147 (estrogen-responsive finger protein)	BG231641
AP-hu02	O	3	O3	2067693	p53-regulated apoptosis-inducing protein 1	AI383639
AP-hu02	O	4	O4	3073056	0	BG231641
AP-hu02	O	5	O5	2090446	polymerase (DNA directed), eta	AI554887
AP-hu02	O	6	O6	3078407	TGFB1-induced anti-apoptotic factor 1	BG231641
AP-hu02	O	7	O7	2094946	RNA binding motif protein 5	AI424264
AP-hu02	O	8	O8	3083083	0	BG231641
AP-hu02	O	9	O9	2108826	interleukin 18 (interferon-gamma-inducing factor)	AI394321
AP-hu02	O	10	O10	3085549	0	BG231641
AP-hu02	O	11	O11	2109792	guanine nucleotide binding protein (G protein), alpha transducing activity polypeptide 2	AF493909
AP-hu02	O	12	O12	3085866	0	BG231641
AP-hu02	O	13	O13	2125556	H2.0-like homeo box 1 (Drosophila)	AI468772
AP-hu02	O	14	O14	3087768	early growth response 2 (Krox-20 homolog, Drosophila)	BG231641
AP-hu02	O	15	O15	2130480	phosphoenolpyruvate carboxykinase 1 (soluble)	AI494177
AP-hu02	O	16	O16	3643475	v-abl Abelson murine leukemia viral oncogene homolog 2 (arg, Abelson-related gene)	BG231641
AP-hu02	O	17	O17	2131329	programmed cell death 8 (apoptosis-inducing factor)	AI431710
AP-hu02	O	18	O18	4142707	killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 2	BG231641

AP-hu02	O	19	O19	2139463	choline/ethanolaminephosphotransferase	AI470615
AP-hu02	O	20	O20	4142901	0	BG231641
AP-hu02	O	21	O21	2144779	WNT1 inducible signaling pathway protein 3	AI457678
AP-hu02	O	22	O22	469577	ribonuclease, RNase A family, 4	AA027189
AP-hu02	O	23	O23	2148030	Bcl-2-associated transcription factor	AI476761
AP-hu02	O	24	O24	469931	TAF9 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 32kDa	AA029766
AP-hu02	P	1	P1	815040	signal peptidase complex (18kD)	AA465269
AP-hu02	P	2	P2	1686798	sushi-repeat protein	AI089690
AP-hu02	P	3	P3	815049	pyruvate dehydrogenase (lipoamide) beta	AA465285
AP-hu02	P	4	P4	1688907	laminin, alpha 4	AI140443
AP-hu02	P	5	P5	898925	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta	AA480228
AP-hu02	P	6	P6	1693375	homeo box A2	AI140882
AP-hu02	P	7	P7	926309	neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, Watson disease)	AA534609
AP-hu02	P	8	P8	1698339	collagen, type XI, alpha 1	AI096671
AP-hu02	P	9	P9	966687	zinc finger protein, subfamily 1A, 5 (Pegasus)	AA505960
AP-hu02	P	10	P10	1702744	heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)	AI096954
AP-hu02	P	11	P11	1030258	Gem-interacting protein	AA554534
AP-hu02	P	12	P12	1702957	NICE-4 protein	AI097419
AP-hu02	P	13	P13	1030484	translocated promoter region (to activated MET oncogene)	AA555031
AP-hu02	P	14	P14	1706935	prefoldin 5	AI126765
AP-hu02	P	15	P15	1031248	solute carrier family 30 (zinc transporter), member 4	AA609034
AP-hu02	P	16	P16	1709428	Fc fragment of IgG, low affinity IIIa, receptor for (CD16)	AI130777
AP-hu02	P	17	P17	1032632	POU domain, class 4, transcription factor 1	AA779554
AP-hu02	P	18	P18	1710655	neuropilin 2	AI139365
AP-hu02	P	19	P19	1032647	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 4	AA779570
AP-hu02	P	20	P20	1712192	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	AI125912
AP-hu02	P	21	P21	1056449	Notch homolog 3 (Drosophila)	AA557201
AP-hu02	P	22	P22	1715851	hemoglobin, gamma G	AI151352
AP-hu02	P	23	P23	1057103	human immunodeficiency virus type I enhancer binding protein 2	AA557381
AP-hu02	P	24	P24	1725451	leukocyte-associated Ig-like receptor 2	AI189606
AP-hu03	A	1	A1	PK_A1	Human 90Kda heat shock protein	#N/A
AP-hu03	A	2	A2	1737502	trinucleotide repeat containing 11 (THR-associated protein, 230kDa subunit)	AI138726
AP-hu03	A	3	A3	PK_A2	60s Acidic Ribosomal Protein PO	#N/A
AP-hu03	A	4	A4	1740083	thyroid transcription factor 1	AI187322
AP-hu03	A	5	A5	PK_A4	nucleophosmin	#N/A
AP-hu03	A	6	A6	1740578	host cell factor C1 (VP16-accessory protein)	AI185261
AP-hu03	A	7	A7	PK_A6	heat shock cognate 70kDa protein	#N/A
AP-hu03	A	8	A8	1756042	modulator of apoptosis 1	AI203801
AP-hu03	A	9	A9	PK_A10	NADH-Ubiquinone oxidoreductase SGDH subunit precusor	#N/A
AP-hu03	A	10	A10	1759710	T-cell leukemia/lymphoma 6	AI219070
AP-hu03	A	11	A11	PK_A11	Human HepG2 3' region cDNA	#N/A
AP-hu03	A	12	A12	1840595	homeo box A5	AI223317
AP-hu03	A	13	A13	PK_B1	Human triosephosphate isomerase	#N/A
AP-hu03	A	14	A14	1841164	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	AI219647
AP-hu03	A	15	A15	PK_B2	Human 3-hydroxy-3-methylglutayl coenzyme A reductase	#N/A
AP-hu03	A	16	A16	1846264	myoglobin	AI239419
AP-hu03	A	17	A17	PK_B3	Elongation factor 1 alpha subunit	#N/A
AP-hu03	A	18	A18	1855723	ret proto-oncogene (multiple endocrine neoplasia and medullary thyroid carcinoma 1, Hirschsprung disease)	AI263439
AP-hu03	A	19	A19	PK_B4	Stratagene fetal retina	#N/A
AP-hu03	A	20	A20	1857711	Wilms tumor 1	AI246677
AP-hu03	A	21	A21	PK_B5	pre mRNA splicing factor SRP20	#N/A
AP-hu03	A	22	A22	1858266	postsynaptic protein CRIPT	AI223792
AP-hu03	A	23	A23	PK_B6	Human acidic ribosomal phosphprotein PO (60s)	#N/A
AP-hu03	A	24	A24	1868332	cell growth regulatory with EF-hand domain	AI261384
AP-hu03	B	1	B1	0	0	#N/A
AP-hu03	B	2	B2	0	0	#N/A
AP-hu03	B	3	B3	0	0	#N/A
AP-hu03	B	4	B4	0	0	#N/A
AP-hu03	B	5	B5	0	0	#N/A
AP-hu03	B	6	B6	0	0	#N/A
AP-hu03	B	7	B7	0	0	#N/A
AP-hu03	B	8	B8	0	0	#N/A
AP-hu03	B	9	B9	0	0	#N/A
AP-hu03	B	10	B10	0	0	#N/A
AP-hu03	B	11	B11	0	0	#N/A
AP-hu03	B	12	B12	0	0	#N/A
AP-hu03	B	13	B13	0	0	#N/A
AP-hu03	B	14	B14	0	0	#N/A
AP-hu03	B	15	B15	0	0	#N/A
AP-hu03	B	16	B16	0	0	#N/A
AP-hu03	B	17	B17	0	0	#N/A
AP-hu03	B	18	B18	0	0	#N/A
AP-hu03	B	19	B19	0	0	#N/A
AP-hu03	B	20	B20	0	0	#N/A
AP-hu03	B	21	B21	0	0	#N/A
AP-hu03	B	22	B22	0	0	#N/A
AP-hu03	B	23	B23	0	0	#N/A
AP-hu03	B	24	B24	0	0	#N/A
AP-hu03	C	1	C1	PK_B7	EST 27474 (Brain-human) and alpha-fetoprotein enhancer binding protein	#N/A
AP-hu03	C	2	C2	1874113	microtubule-associated protein 4	AI339709
AP-hu03	C	3	C3	PK_B8	Human ribosomal protein L30	#N/A
AP-hu03	C	4	C4	1881682	GS2 gene	AI290788
AP-hu03	C	5	C5	PK_B9	Peptidyl-prolyl cis-trans isomerase (PPIASE) Rotamase	#N/A
AP-hu03	C	6	C6	1926660	heterogeneous nuclear ribonucleoprotein H1 (H)	AI346731
AP-hu03	C	7	C7	PK_B10	Ubiquinol-cytochrome-C reductase complex core protein1	#N/A

AP-hu03	C	8	C8	1929994	fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)	AI313221
AP-hu03	C	9	C9	PK_B11	3 LINE 1 elements	#N/A
AP-hu03	C	10	C10	1949396	thromboxane A synthase 1 (platelet, cytochrome P450, family 5, subfamily A)	AI342602
AP-hu03	C	11	C11	PK_B12	Trans-acting transcriptional protein ICPO	#N/A
AP-hu03	C	12	C12	2014456	prominin 1	AI362100
AP-hu03	C	13	C13	PK_C1	Human tumour necrosis factor type 1 receptor associated protein	#N/A
AP-hu03	C	14	C14	2015313	adhesion glycoprotein	AI361991
AP-hu03	C	15	C15	PK_C3	peptidyl-prolyl:cis-trans isomerase A	#N/A
AP-hu03	C	16	C16	2015920	natural killer cell receptor 2B4	AI363168
AP-hu03	C	17	C17	PK_C4	PK_C4	#N/A
AP-hu03	C	18	C18	2029834	general transcription factor IIH, polypeptide 3, 34kDa	AI373329
AP-hu03	C	19	C19	PK_C5	Eukaryotic initiation factor 4B (EIF-4B)	#N/A
AP-hu03	C	20	C20	2042756	laminin, alpha 3	AI375787
AP-hu03	C	21	C21	PK_C6	PK_C6	#N/A
AP-hu03	C	22	C22	2063257	0	AI382794
AP-hu03	C	23	C23	PK_C8	Guanine nucleotide-binding protein beta subunit-like protein	#N/A
AP-hu03	C	24	C24	2066379	chloride channel 4	AI802379
AP-hu03	D	1	D1	0	0	#N/A
AP-hu03	D	2	D2	0	0	#N/A
AP-hu03	D	3	D3	0	0	#N/A
AP-hu03	D	4	D4	0	0	#N/A
AP-hu03	D	5	D5	0	0	#N/A
AP-hu03	D	6	D6	0	0	#N/A
AP-hu03	D	7	D7	0	0	#N/A
AP-hu03	D	8	D8	0	0	#N/A
AP-hu03	D	9	D9	0	0	#N/A
AP-hu03	D	10	D10	0	0	#N/A
AP-hu03	D	11	D11	0	0	#N/A
AP-hu03	D	12	D12	0	0	#N/A
AP-hu03	D	13	D13	0	0	#N/A
AP-hu03	D	14	D14	0	0	#N/A
AP-hu03	D	15	D15	0	0	#N/A
AP-hu03	D	16	D16	0	0	#N/A
AP-hu03	D	17	D17	0	0	#N/A
AP-hu03	D	18	D18	0	0	#N/A
AP-hu03	D	19	D19	0	0	#N/A
AP-hu03	D	20	D20	0	0	#N/A
AP-hu03	D	21	D21	0	0	#N/A
AP-hu03	D	22	D22	0	0	#N/A
AP-hu03	D	23	D23	0	0	#N/A
AP-hu03	D	24	D24	0	0	#N/A
AP-hu03	E	1	E1	PK_C9	elongation factor (yeast)	#N/A
AP-hu03	E	2	E2	2067266	histone deacetylase 8	AI803366
AP-hu03	E	3	E3	PK_C10	Guanine nucleotide-binding protein beta subunit-like protein	#N/A
AP-hu03	E	4	E4	2068299	adrenomedullin receptor	AI379917
AP-hu03	E	5	E5	PK_C12	Soares pregant uterur NbHPU Homo sapien cDNA clone	#N/A
AP-hu03	E	6	E6	2068829	joined to JAZF1	AI378171
AP-hu03	E	7	E7	PK_D2	Yeast Pep II protein	#N/A
AP-hu03	E	8	E8	2072912	apoptosis antagonizing transcription factor	AI439571
AP-hu03	E	9	E9	PK_D3	human Pep II protein	#N/A
AP-hu03	E	10	E10	2091078	ras homolog gene family, member N	AI377326
AP-hu03	E	11	E11	PK_D5	PK_D5	#N/A
AP-hu03	E	12	E12	2094592	activin A receptor type II-like 1	AI420822
AP-hu03	E	13	E13	PK_D8	16s rRNA	#N/A
AP-hu03	E	14	E14	2095272	regulator of G-protein signalling 9	AI420112
AP-hu03	E	15	E15	PK_D10	Heat shock cognate 71kDa protein	#N/A
AP-hu03	E	16	E16	2096365	tumor protein p53 inducible nuclear protein 1	AI419377
AP-hu03	E	17	E17	PK_D11	PK_D11	#N/A
AP-hu03	E	18	E18	2098563	transcription elongation factor A (SII), 1	AI421762
AP-hu03	E	19	E19	PK_E1	HLA class II histocompatiblity antigen	#N/A
AP-hu03	E	20	E20	2108356	suppressor of Ty 6 homolog (S. cerevisiae)	AI392835
AP-hu03	E	21	E21	PK_E3	Soares parathyroid tumour NbHPA	#N/A
AP-hu03	E	22	E22	2108870	chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	AI391519
AP-hu03	E	23	E23	PK_E6	Human ribosomal DNA	#N/A
AP-hu03	E	24	E24	2111581	mitogen-activated protein kinase 11	AI394426
AP-hu03	F	1	F1	0	0	#N/A
AP-hu03	F	2	F2	0	0	#N/A
AP-hu03	F	3	F3	0	0	#N/A
AP-hu03	F	4	F4	0	0	#N/A
AP-hu03	F	5	F5	0	0	#N/A
AP-hu03	F	6	F6	0	0	#N/A
AP-hu03	F	7	F7	0	0	#N/A
AP-hu03	F	8	F8	0	0	#N/A
AP-hu03	F	9	F9	0	0	#N/A
AP-hu03	F	10	F10	0	0	#N/A
AP-hu03	F	11	F11	0	0	#N/A
AP-hu03	F	12	F12	0	0	#N/A
AP-hu03	F	13	F13	0	0	#N/A
AP-hu03	F	14	F14	0	0	#N/A
AP-hu03	F	15	F15	0	0	#N/A
AP-hu03	F	16	F16	0	0	#N/A
AP-hu03	F	17	F17	0	0	#N/A
AP-hu03	F	18	F18	0	0	#N/A
AP-hu03	F	19	F19	0	0	#N/A
AP-hu03	F	20	F20	0	0	#N/A
AP-hu03	F	21	F21	0	0	#N/A
AP-hu03	F	22	F22	0	0	#N/A
AP-hu03	F	23	F23	0	0	#N/A

AP-hu03	F	24	F24	0	0	#N/A
AP-hu03	G	1	G1	PK_E9	Tubulin gaunachain	#N/A
AP-hu03	G	2	G2	2113938	amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like)	AI400246
AP-hu03	G	3	G3	PK_E11	alu subfamily	#N/A
AP-hu03	G	4	G4	2122503	leukotriene A4 hydrolase	AI521212
AP-hu03	G	5	G5	PK_E12	Human placenta cDNA 5'	#N/A
AP-hu03	G	6	G6	2137887	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	AI538195
AP-hu03	G	7	G7	PK_F1	Colon Carcinoma HCC	#N/A
AP-hu03	G	8	G8	2138953	general transcription factor IIH, polypeptide 1, 62kDa	AI439100
AP-hu03	G	9	G9	PK_F3	60s ribosomal protein L30	#N/A
AP-hu03	G	10	G10	2148566	0	AI469873
AP-hu03	G	11	G11	PK_F6	PK_F6	#N/A
AP-hu03	G	12	G12	2149893	Homo sapiens transcribed sequence with moderate similarity to protein sp:Q92915 (H.sapiens) FGFE_HUMAN Fibroblast growth factor-14 (FGF-14) (Fibroblast growth factor homologous factor 4) (FHF-4)	AI887836
AP-hu03	G	13	G13	PK_F8	40s ribosomal protein s18	#N/A
AP-hu03	G	14	G14	2151139	programmed cell death 8 (apoptosis-inducing factor)	AI589001
AP-hu03	G	15	G15	PK_F10	CAMP-regulated phosphoprotein	#N/A
AP-hu03	G	16	G16	2214059	isocitrate dehydrogenase 3 (NAD+) beta	AI566177
AP-hu03	G	17	G17	PK_F11	40s ribosomal protein s3A	#N/A
AP-hu03	G	18	G18	2237098	low density lipoprotein receptor-related protein 6	AI633264
AP-hu03	G	19	G19	empty	0	#N/A
AP-hu03	G	20	G20	2241908	thousand and one amino acid protein kinase	AI637830
AP-hu03	G	21	G21	empty	0	#N/A
AP-hu03	G	22	G22	2242601	pyruvate carboxylase	AI637857
AP-hu03	G	23	G23	empty	0	#N/A
AP-hu03	G	24	G24	2243192	transcriptional adaptor 2 (ADA2 homolog, yeast)-like	AI655771
AP-hu03	H	1	H1	0	0	#N/A
AP-hu03	H	2	H2	0	0	#N/A
AP-hu03	H	3	H3	0	0	#N/A
AP-hu03	H	4	H4	0	0	#N/A
AP-hu03	H	5	H5	0	0	#N/A
AP-hu03	H	6	H6	0	0	#N/A
AP-hu03	H	7	H7	0	0	#N/A
AP-hu03	H	8	H8	0	0	#N/A
AP-hu03	H	9	H9	0	0	#N/A
AP-hu03	H	10	H10	0	0	#N/A
AP-hu03	H	11	H11	0	0	#N/A
AP-hu03	H	12	H12	0	0	#N/A
AP-hu03	H	13	H13	0	0	#N/A
AP-hu03	H	14	H14	0	0	#N/A
AP-hu03	H	15	H15	0	0	#N/A
AP-hu03	H	16	H16	0	0	#N/A
AP-hu03	H	17	H17	0	0	#N/A
AP-hu03	H	18	H18	0	0	#N/A
AP-hu03	H	19	H19	0	0	#N/A
AP-hu03	H	20	H20	0	0	#N/A
AP-hu03	H	21	H21	0	0	#N/A
AP-hu03	H	22	H22	0	0	#N/A
AP-hu03	H	23	H23	0	0	#N/A
AP-hu03	H	24	H24	0	0	#N/A
AP-hu03	I	1	I1	empty	0	#N/A
AP-hu03	I	2	I2	2243934	interleukin 21 receptor	AI656754
AP-hu03	I	3	I3	empty	0	#N/A
AP-hu03	I	4	I4	2244213	E3-binding protein	AI657202
AP-hu03	I	5	I5	empty	0	#N/A
AP-hu03	I	6	I6	2251783	potassium voltage-gated channel, shaker-related subfamily, beta member 1	AI658686
AP-hu03	I	7	I7	empty	0	#N/A
AP-hu03	I	8	I8	2252681	chemokine (C motif) ligand 1	AI669817
AP-hu03	I	9	I9	empty	0	#N/A
AP-hu03	I	10	I10	2273173	wingless-type MMTV integration site family, member 4	AI634496
AP-hu03	I	11	I11	empty	0	#N/A
AP-hu03	I	12	I12	2284066	vanin 1	AI625852
AP-hu03	I	13	I13	empty	0	#N/A
AP-hu03	I	14	I14	2284081	stanniocalcin 1	AI625879
AP-hu03	I	15	I15	empty	0	#N/A
AP-hu03	I	16	I16	2284473	zinc finger protein 443	AI613104
AP-hu03	I	17	I17	empty	0	#N/A
AP-hu03	I	18	I18	2304672	telomerase reverse transcriptase	AI824948
AP-hu03	I	19	I19	empty	0	#N/A
AP-hu03	I	20	I20	2305123	Homo sapiens transcribed sequences	AI631732
AP-hu03	I	21	I21	empty	0	#N/A
AP-hu03	I	22	I22	2316826	cytochrome P450, family 3, subfamily A, polypeptide 4	AI671993
AP-hu03	I	23	I23	empty	0	#N/A
AP-hu03	I	24	I24	2321108	basic leucine zipper nuclear factor 1 (JEM-1)	AI675660
AP-hu03	J	1	J1	0	0	#N/A
AP-hu03	J	2	J2	0	0	#N/A
AP-hu03	J	3	J3	0	0	#N/A
AP-hu03	J	4	J4	0	0	#N/A
AP-hu03	J	5	J5	0	0	#N/A
AP-hu03	J	6	J6	0	0	#N/A
AP-hu03	J	7	J7	0	0	#N/A
AP-hu03	J	8	J8	0	0	#N/A
AP-hu03	J	9	J9	0	0	#N/A
AP-hu03	J	10	J10	0	0	#N/A
AP-hu03	J	11	J11	0	0	#N/A
AP-hu03	J	12	J12	0	0	#N/A
AP-hu03	J	13	J13	0	0	#N/A

AP-hu03	J	14	J14	0	0	#N/A
AP-hu03	J	15	J15	0	0	#N/A
AP-hu03	J	16	J16	0	0	#N/A
AP-hu03	J	17	J17	0	0	#N/A
AP-hu03	J	18	J18	0	0	#N/A
AP-hu03	J	19	J19	0	0	#N/A
AP-hu03	J	20	J20	0	0	#N/A
AP-hu03	J	21	J21	0	0	#N/A
AP-hu03	J	22	J22	0	0	#N/A
AP-hu03	J	23	J23	0	0	#N/A
AP-hu03	J	24	J24	0	0	#N/A
AP-hu03	K	1	K1	empty	0	#N/A
AP-hu03	K	2	K2	2330252	RAD51-like 3 (S. cerevisiae)	AI692982
AP-hu03	K	3	K3	empty	0	#N/A
AP-hu03	K	4	K4	3068365	transient receptor potential cation channel, subfamily V, member 6	AW452397
AP-hu03	K	5	K5	empty	0	#N/A
AP-hu03	K	6	K6	3070183	RNA binding motif protein 5	BF513550
AP-hu03	K	7	K7	empty	0	#N/A
AP-hu03	K	8	K8	3071696	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3	BF513730
AP-hu03	K	9	K9	empty	0	#N/A
AP-hu03	K	10	K10	2731796	TAF4b RNA polymerase II, TATA box binding protein (TBP)-associated factor, 105kDa	AW297262
AP-hu03	K	11	K11	empty	0	#N/A
AP-hu03	K	12	K12	2731913	Homo sapiens transcribed sequences	AW297706
AP-hu03	K	13	K13	empty	0	#N/A
AP-hu03	K	14	K14	3173557	regulator of G-protein signalling 4	BE222672
AP-hu03	K	15	K15	empty	0	#N/A
AP-hu03	K	16	K16	3178227	programmed cell death 6	BE219939
AP-hu03	K	17	K17	empty	0	#N/A
AP-hu03	K	18	K18	3183031	lipoprotein lipase	BE465407
AP-hu03	K	19	K19	empty	0	#N/A
AP-hu03	K	20	K20	3184548	epidermal growth factor (beta-urogastrone)	BE501422
AP-hu03	K	21	K21	empty	0	#N/A
AP-hu03	K	22	K22	3196945	aminoadipate-semialdehyde synthase	BE466144
AP-hu03	K	23	K23	empty	0	#N/A
AP-hu03	K	24	K24	3209795	adrenergic, beta, receptor kinase 2	BE504864
AP-hu03	L	1	L1	0	0	#N/A
AP-hu03	L	2	L2	0	0	#N/A
AP-hu03	L	3	L3	0	0	#N/A
AP-hu03	L	4	L4	0	0	#N/A
AP-hu03	L	5	L5	0	0	#N/A
AP-hu03	L	6	L6	0	0	#N/A
AP-hu03	L	7	L7	0	0	#N/A
AP-hu03	L	8	L8	0	0	#N/A
AP-hu03	L	9	L9	0	0	#N/A
AP-hu03	L	10	L10	0	0	#N/A
AP-hu03	L	11	L11	0	0	#N/A
AP-hu03	L	12	L12	0	0	#N/A
AP-hu03	L	13	L13	0	0	#N/A
AP-hu03	L	14	L14	0	0	#N/A
AP-hu03	L	15	L15	0	0	#N/A
AP-hu03	L	16	L16	0	0	#N/A
AP-hu03	L	17	L17	0	0	#N/A
AP-hu03	L	18	L18	0	0	#N/A
AP-hu03	L	19	L19	0	0	#N/A
AP-hu03	L	20	L20	0	0	#N/A
AP-hu03	L	21	L21	0	0	#N/A
AP-hu03	L	22	L22	0	0	#N/A
AP-hu03	L	23	L23	0	0	#N/A
AP-hu03	L	24	L24	0	0	#N/A
AP-hu03	M	1	M1	empty	0	#N/A
AP-hu03	M	2	M2	3213061	19A24 protein	BE467670
AP-hu03	M	3	M3	empty	0	#N/A
AP-hu03	M	4	M4	3218773	hypothetical protein FLJ20793	BE503286
AP-hu03	M	5	M5	empty	0	#N/A
AP-hu03	M	6	M6	3220173	Cas-Br-M (murine) ecotropic retroviral transforming sequence	BE550582
AP-hu03	M	7	M7	empty	0	#N/A
AP-hu03	M	8	M8	3220758	Homo sapiens transcribed sequences	BE673055
AP-hu03	M	9	M9	empty	0	#N/A
AP-hu03	M	10	M10	3230943	insulinoma-associated 1	BE549911
AP-hu03	M	11	M11	empty	0	#N/A
AP-hu03	M	12	M12	3286412	complement component 1, q subcomponent, receptor 1	BE671226
AP-hu03	M	13	M13	empty	0	#N/A
AP-hu03	M	14	M14	3316202	protein tyrosine phosphatase, receptor type, H	BF000072
AP-hu03	M	15	M15	empty	0	#N/A
AP-hu03	M	16	M16	3316686	glycogen synthase 1 (muscle)	BF002321
AP-hu03	M	17	M17	empty	0	#N/A
AP-hu03	M	18	M18	3318799	MARCKS-like protein	BF591250
AP-hu03	M	19	M19	empty	0	#N/A
AP-hu03	M	20	M20	3320058	cytochrome P450, family 3, subfamily A, polypeptide 5	BF591425
AP-hu03	M	21	M21	empty	0	#N/A
AP-hu03	M	22	M22	3393054	retinoblastoma binding protein 5	BF055070
AP-hu03	M	23	M23	empty	0	#N/A
AP-hu03	M	24	M24	3405658	transforming growth factor, beta receptor I (activin A receptor type II-like kinase, 53kDa)	BF439832
AP-hu03	N	1	N1	0	0	#N/A
AP-hu03	N	2	N2	0	0	#N/A
AP-hu03	N	3	N3	0	0	#N/A

AP-hu03	N	4	N4	0	0	#N/A
AP-hu03	N	5	N5	0	0	#N/A
AP-hu03	N	6	N6	0	0	#N/A
AP-hu03	N	7	N7	0	0	#N/A
AP-hu03	N	8	N8	0	0	#N/A
AP-hu03	N	9	N9	0	0	#N/A
AP-hu03	N	10	N10	0	0	#N/A
AP-hu03	N	11	N11	0	0	#N/A
AP-hu03	N	12	N12	0	0	#N/A
AP-hu03	N	13	N13	0	0	#N/A
AP-hu03	N	14	N14	0	0	#N/A
AP-hu03	N	15	N15	0	0	#N/A
AP-hu03	N	16	N16	0	0	#N/A
AP-hu03	N	17	N17	0	0	#N/A
AP-hu03	N	18	N18	0	0	#N/A
AP-hu03	N	19	N19	0	0	#N/A
AP-hu03	N	20	N20	0	0	#N/A
AP-hu03	N	21	N21	0	0	#N/A
AP-hu03	N	22	N22	0	0	#N/A
AP-hu03	N	23	N23	0	0	#N/A
AP-hu03	N	24	N24	0	0	#N/A
AP-hu03	O	1	O1	empty	0	#N/A
AP-hu03	O	2	O2	3476767	arylsulfatase B	BF058213
AP-hu03	O	3	O3	empty	0	#N/A
AP-hu03	O	4	O4	3477610	Wilms tumor associated protein	BF059076
AP-hu03	O	5	O5	empty	0	#N/A
AP-hu03	O	6	O6	3523521	integrin, beta 6	BF111900
AP-hu03	O	7	O7	empty	0	#N/A
AP-hu03	O	8	O8	3525740	protein transport protein SEC61 alpha subunit isoform 1	BF109012
AP-hu03	O	9	O9	empty	0	#N/A
AP-hu03	O	10	O10	3562165	postmeiotic segregation increased 2-like 6	BF197176
AP-hu03	O	11	O11	empty	0	#N/A
AP-hu03	O	12	O12	3562195	YY1 associated factor 2	BF197072
AP-hu03	O	13	O13	empty	0	#N/A
AP-hu03	O	14	O14	3085383	zinc finger protein 256	BF510334
AP-hu03	O	15	O15	empty	0	#N/A
AP-hu03	O	16	O16	3085523	spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration)	BF510366
AP-hu03	O	17	O17	empty	0	#N/A
AP-hu03	O	18	O18	3565961	chemokine (C-C motif) ligand 22	BF432854
AP-hu03	O	19	O19	empty	0	#N/A
AP-hu03	O	20	O20	3570409	B cell RAG associated protein	BF115957
AP-hu03	O	21	O21	empty	0	#N/A
AP-hu03	O	22	O22	empty	0	#N/A
AP-hu03	O	23	O23	empty	0	#N/A
AP-hu03	O	24	O24	empty	0	#N/A
AP-hu03	P	1	P1	0	0	#N/A
AP-hu03	P	2	P2	0	0	#N/A
AP-hu03	P	3	P3	0	0	#N/A
AP-hu03	P	4	P4	0	0	#N/A
AP-hu03	P	5	P5	0	0	#N/A
AP-hu03	P	6	P6	0	0	#N/A
AP-hu03	P	7	P7	0	0	#N/A
AP-hu03	P	8	P8	0	0	#N/A
AP-hu03	P	9	P9	0	0	#N/A
AP-hu03	P	10	P10	0	0	#N/A
AP-hu03	P	11	P11	0	0	#N/A
AP-hu03	P	12	P12	0	0	#N/A
AP-hu03	P	13	P13	0	0	#N/A
AP-hu03	P	14	P14	0	0	#N/A
AP-hu03	P	15	P15	0	0	#N/A
AP-hu03	P	16	P16	0	0	#N/A
AP-hu03	P	17	P17	0	0	#N/A
AP-hu03	P	18	P18	0	0	#N/A
AP-hu03	P	19	P19	0	0	#N/A
AP-hu03	P	20	P20	0	0	#N/A
AP-hu03	P	21	P21	0	0	#N/A
AP-hu03	P	22	P22	0	0	#N/A
AP-hu03	P	23	P23	0	0	#N/A
AP-hu03	P	24	P24	0	0	#N/A

Appendix C - Sequence alignments of RT-PCR amplified splice variants of HHV-7

>U17 top band

```
EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 650 (103.6 bits), Expect = 3.9e-21, P = 3.9e-21
 Identities = 208/280 (74%), Positives = 208/280 (74%), Strand = Plus / Plus

Query:     2 TTGCA-CATAGGAGATAGCAACCTATGAATTGA-CAAACACTTTCTATAACATCTTGACA 59
             ||||| |||||||||||||||||||||||||   | |||| ||||  || |||||| |||||
Sbjct: 27747 TTGCAACATAGGAGATAGCAACCTATGAAAACAACAAAAACTTATTAGAACATCATGACA 27806

Query:    60 CGGTGGAAATGGTTTGACTGCTTTTGTA-CTT-CATACTCTCTTGCCA-TCT-TATGAAG 115
             |  |||||||| ||||||||||||||| | ||| ||||   ||||||   ||  | | |
Sbjct: 27807 CATTGGAAATGTTTTGACTGCTTTTGCAACTTACATAGG-TCTTGCAGCTCGGTTTCACT 27865

Query:   116 ATTTCACACCTGATATGTTGATCCACG-AAAGGCCTAATAAAATTTTTCTTAAT--G-GG 171
             | |  |   |||| | |   |||| || ||||||  ||| |   ||||| || |  | ||
Sbjct: 27866 AGT--AGTTCTGAAAGGCATATCC-CGCAAAGGCGAAATGATTATTTTCCTACTCTGAGG 27922

Query:   172 TTGTCTAACAAGGATTAAATTTCCCAGTGTATGCCTAATAACGAAGTCACTAAGAGAATC 231
               ||| |  | ||||  |   |||| | | |    | ||||||||||||||||||||
Sbjct: 27923 CCATCTTATGAAGATTTCACA-CCCACTATTTTG--AGTAACGAAGTCACTAAGAGAATC 27979

Query:   232 GCGATAAAAGGCTTGTTTAACAAGGCTTAAATTTGCGAGT 271
             |||||||||||||||||||||||||||||||||||||||
Sbjct: 27980 GCGATAAAAGGCTTGTTTAACAAGGCTTAAATTTGCGAGT 28019


 Score = 349 (58.4 bits), Expect = 1.5e-07, P = 1.5e-07
 Identities = 99/131 (75%), Positives = 99/131 (75%), Strand = Plus / Plus

Query:    89 TTCATACTCTCTTGCCATCTTATGAAGATTTCACACCTGATATGTTGATCCACGAAAGGC 148
             ||| ||||||   |||||||||||||||||||||||   ||| ||||   ||||| | |
Sbjct: 27909 TTCCTACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATTTTGAGTAACGAA-GTC 27967

Query:   149 -CTAATAAAATTTTTCTTAATGGGTTGTCTAACAAGGATTAAATTTCCCAGTGTATGCCT 207
              |||| | |||     | || || |||| |||||||| |||||||| | |||| ||||||
Sbjct: 27968 ACTAAGAGAATCGCGATAAAAGGCTTGTTTAACAAGGCTTAAATTTGCGAGTGCATGCCT 28027

Query:   208 AATAACGAAGT 218
             || | |||| |
Sbjct: 28028 AA-ATCGAACT 28037
```

>U17 mid band

```
>EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 1010 (157.6 bits), Expect = 2.1e-37, P = 2.1e-37
 Identities = 202/202 (100%), Positives = 202/202 (100%), Strand = Plus / Plus

Query:     1 AGAACATCATGACACATTGGAAATGTTTTGACTGCTTTTGCAACTTACATAGGTCTTGCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27793 AGAACATCATGACACATTGGAAATGTTTTGACTGCTTTTGCAACTTACATAGGTCTTGCA 27852

Query:    61 GCTCGGTTTCACTAGTAGTTCTGAAAGGCATATCCCGCAAAGGCGAAATGATTATTTTCC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27853 GCTCGGTTTCACTAGTAGTTCTGAAAGGCATATCCCGCAAAGGCGAAATGATTATTTTCC 27912

Query:   121 TACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATTTTGAGTAACGAAGTCACTAA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27913 TACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATTTTGAGTAACGAAGTCACTAA 27972

Query:   181 GAGAATCGCGATAAAAGGCTTG 202
             |||||||||||||||||||||
Sbjct: 27973 GAGAATCGCGATAAAAGGCTTG 27994
```

>U17 bottom band

```
EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 870 (136.6 bits), Expect = 4.5e-31, P = 4.5e-31
 Identities = 174/174 (100%), Positives = 174/174 (100%), Strand = Plus / Plus

Query:     15 ACATAGGTCTTGCAGCTCGGTTTCACTAGTAGTTCTGAAAGGCATATCCCGCAAAGGCGA 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27839 ACATAGGTCTTGCAGCTCGGTTTCACTAGTAGTTCTGAAAGGCATATCCCGCAAAGGCGA 27898

Query:     75 AATGATTATTTTCCTACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATTTTGAGT 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27899 AATGATTATTTTCCTACTCTGAGGCCATCTTATGAAGATTTCACACCCACTATTTTGAGT 27958

Query:    135 AACGAAGTCACTAAGAGAATCGCGATAAAAGGCTTGTTTAACAAGGCTTAAATT 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27959 AACGAAGTCACTAAGAGAATCGCGATAAAAGGCTTGTTTAACAAGGCTTAAATT 28012
```

>U90

```
EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 418 (68.8 bits), Expect = 1.2e-10, P = 1.2e-10
 Identities = 90/95 (94%), Positives = 90/95 (94%), Strand = Plus / Plus

Query:      1 AGAGTATTATCAAAGGAAGGGTTGCATTTTAAACTGTCTTGCAATATGTCTCTGATGGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 133158 AGAGTATTATCAAAGGAAGGGTTGCATTTTAAACTGTCTTGCAATATGTCTCTGATGGTC 133217

Query:     61 TCTTCCACGGGGGTATCTGGGCGATAT-TC-AATA 93
              |||||||||| |||||||||| ||||| |  ||||
Sbjct: 133218 TCTTCCACGGTGGTATCTGGGTGATATCTGTAATA 133252
```

>U41 top band

```
>EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 614 (98.2 bits), Expect = 1.7e-19, P = 1.7e-19
 Identities = 136/151 (90%), Positives = 136/151 (90%), Strand = Plus / Plus

Query:      1 GATTTTGGTAATGACTGTTCCGCCGAAATTAATAATGGGTGTTTTAACGGTTGGGAAGAA 60
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct: 66386 GATTTTGGTAATGACTGTTCCGCCGTAATTAATAATGGGTGTTTTAACGGTTGGGAAGAA 66445

Query:     61 GTCATTTTCACCCAACCGATTAAGTAACACGGGTGAGATCCCCTCCGATTTTCTTTTTTC 120
              |||||||||| ||    |||||||||||| |||||||||| |  | ||||||||||||||
Sbjct: 66446 GTCATTTTCAACCGTTAGATTAAGTAACAAGGGTGAGATCACTACTGATTTTCTTTTTTC 66505

Query:    121 CATTAACGATAATGTTGAAAGAATTTCATAT 151
              ||||||||||||||| |  |||||||| |||
Sbjct: 66506 CATTAACGATAATGTAGTCAGAATTTC-TAT 66535
```

>CL6 = U41 bottom band

```
EM_VI:AF037218 AF037218.1 Human herpesvirus 7 strain RK, complete genome.
        Length = 153,080
  Plus Strand HSPs:
Score = 885 (138.8 bits), Expect = 9.5e-32, P = 9.5e-32
 Identities = 177/177 (100%), Positives = 177/177 (100%), Strand = Plus / Plus

Query:      1 ATACAGGCATAAACGATGTGATTTTGGTAATGACTGTTCCGCCGTAATTAATAATGGGTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 66367 ATACAGGCATAAACGATGTGATTTTGGTAATGACTGTTCCGCCGTAATTAATAATGGGTG 66426

Query:     61 TTTTAACGGTTGGGAAGAAGTCATTTTCAACCGTTAGATTAAGTAACAAGGGTGAGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 66427 TTTTAACGGTTGGGAAGAAGTCATTTTCAACCGTTAGATTAAGTAACAAGGGTGAGATCA 66486

Query:    121 CTACTGATTTTCTTTTTTCCATTAACGATAATGTAGTCAGAATTTCTATTAATTTTT 177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 66487 CTACTGATTTTCTTTTTTCCATTAACGATAATGTAGTCAGAATTTCTATTAATTTTT 66543
```

REFERENCES

Ablashi, D. V., Handy, M., Bernbaum, J., Chatlynne, L. G., Lapps, W., Kramarsky, B., Berneman, Z. N., Komaroff, A. L., and Whitman, J. E. (1998). Propagation and characterization of human herpesvirus-7 (HHV-7) isolates in a continuous T-lymphoblastoid cell line (SupT1). J Virol Methods *73*, 123-140.

Agaisse, H., Burrack, L. S., Philips, J. A., Rubin, E. J., Perrimon, N., and Higgins, D. E. (2005). Genome-wide RNAi screen for host factors required for intracellular bacterial infection. Science *309*, 1248-1251.

Ahn, J. H., Jang, W. J., and Hayward, G. S. (1999). The human cytomegalovirus IE2 and UL112-113 proteins accumulate in viral DNA replication compartments that initiate from the periphery of promyelocytic leukemia protein-associated nuclear bodies (PODs or ND10). J Virol *73*, 10458-10471.

Ahn, J. W., Powell, K. L., Kellam, P., and Alber, D. G. (2002). Gammaherpesvirus lytic gene expression as characterized by DNA array. J Virol *76*, 6244-6256.

Alba, M. M., Lee, D., Pearl, F. M., Shepherd, A. J., Martin, N., Orengo, C. A., and Kellam, P. (2001). VIDA: a virus database system for the organization of animal virus genome open reading frames. Nucleic Acids Res *29*, 133-136.

Alsaker, K. V., and Papoutsakis, E. T. (2005). Transcriptional program of early sporulation and stationary-phase events in Clostridium acetobutylicum. J Bacteriol *187*, 7103-7118.

Altschul, S. F., Bundschuh, R., Olsen, R., and Hwa, T. (2001). The estimation of statistical parameters for local alignment score distributions. Nucleic Acids Res *29*, 351-361.

Bahnson, A. B., Dunigan, J. T., Baysal, B. E., Mohney, T., Atchison, R. W., Nimgaonkar, M. T., Ball, E. D., and Barranger, J. A. (1995). Centrifugal enhancement of retroviral mediated gene transfer. J Virol Methods *54*, 131-143.

Baldick, C. J., Jr., Marchini, A., Patterson, C. E., and Shenk, T. (1997). Human cytomegalovirus tegument protein pp71 (ppUL82) enhances the infectivity of viral DNA and accelerates the infectious cycle. J Virol *71*, 4400-4408.

Bassett, D. E., Jr., Eisen, M. B., and Boguski, M. S. (1999). Gene expression informatics--it's all in your mine. Nat Genet *21*, 51-55.

Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S., and Small, P. M. (1999). Comparative genomics of BCG vaccines by whole-genome DNA microarray. Science *284*, 1520-1523.

Ben Mamoun, C., Gluzman, I. Y., Hott, C., MacMillan, S. K., Amarakone, A. S., Anderson, D. L., Carlton, J. M., Dame, J. B., Chakrabarti, D., Martin, R. K., *et al.* (2001). Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite Plasmodium falciparum revealed by microarray analysis. Mol Microbiol *39*, 26-36.

Berneman, Z. N., Ablashi, D. V., Li, G., Eger-Fletcher, M., Reitz, M. S., Jr., Hung, C. L., Brus, I., Komaroff, A. L., and Gallo, R. C. (1992a). Human herpesvirus 7 is a T-lymphotropic virus and is

related to, but significantly different from, human herpesvirus 6 and human cytomegalovirus. Proc Natl Acad Sci U S A *89*, 10552-10556.

Berneman, Z. N., Gallo, R. C., Ablashi, D. V., Frenkel, N., Katsafanas, G., Kramarsky, B., and Brus, I. (1992b). Human herpesvirus 7 (HHV-7) strain JI: independent confirmation of HHV-7. J Infect Dis *166*, 690-691.

Bigger, C. B., Brasky, K. M., and Lanford, R. E. (2001). DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. J Virol *75*, 7059-7066.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res *7*, 1513-1523.

Black, J. B., Burns, D. A., Goldsmith, C. S., Feorino, P. M., Kite-Powell, K., Schinazi, R. F., Krug, P. W., and Pellett, P. E. (1997). Biologic properties of human herpesvirus 7 strain SB. Virus Res *52*, 25-41.

Black, J. B., and Pellett, P. E. (1999). Human herpesvirus 7. Rev Med Virol *9*, 245-262.

Black, J. B., Sanderlin, K. C., Goldsmith, C. S., Gary, H. E., Lopez, C., and Pellett, P. E. (1989). Growth properties of human herpesvirus-6 strain Z29. J Virol Methods *26*, 133-145.

Bolt, G., Berg, K., and Blixenkrone-Moller, M. (2002). Measles virus-induced modulation of host-cell gene expression. J Gen Virol *83*, 1157-1165.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H. C., *et al.* (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet *29*, 365-371.

Brennan, D. C., Storch, G. A., Singer, G. G., Lee, L., Rueda, J., and Schnitzler, M. A. (2000). The prevalence of human herpesvirus-7 in renal transplant recipients is unaffected by oral or intravenous ganciclovir. J Infect Dis *181*, 1557-1561.

Browne, E. P., Wing, B., Coleman, D., and Shenk, T. (2001). Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. J Virol *75*, 12319-12330.

Broyles, S. S. (2003). Vaccinia virus transcription. J Gen Virol *84*, 2293-2303.

Cai, H. Y., Lu, L., Muckle, C. A., Prescott, J. F., and Chen, S. (2005). Development of a novel protein microarray method for serotyping Salmonella enterica strains. J Clin Microbiol *43*, 3427-3430.

Campadelli-Fiume, G., Cocchi, F., Menotti, L., and Lopez, M. (2000). The novel receptors that mediate the entry of herpes simplex viruses and animal alphaherpesviruses into cells. Rev Med Virol *10*, 305-319.

Carter, K. L., Cahir-McFarland, E., and Kieff, E. (2002). Epstein-barr virus-induced changes in B-lymphocyte gene expression. J Virol *76*, 10427-10436.

Chambers, J., Angulo, A., Amaratunga, D., Guo, H., Jiang, Y., Wan, J. S., Bittner, A., Frueh, K., Jackson, M. R., Peterson, P. A., *et al.* (1999). DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. J Virol *73*, 5757-5766.

Chan, P. K., Peiris, J. S., Yuen, K. Y., Liang, R. H., Lau, Y. L., Chen, F. E., Lo, S. K., Cheung, C. Y., Chan, T. K., and Ng, M. H. (1997). Human herpesvirus-6 and human herpesvirus-7 infections in bone marrow transplant recipients. J Med Virol *53*, 295-305.

Chang, Y. E., and Laimins, L. A. (2000). Microarray analysis identifies interferon-inducible genes and Stat-1 as major transcriptional targets of human papillomavirus type 31. J Virol *74*, 4174-4182.

Che, S., and Ginsberg, S. D. (2004). Amplification of RNA transcripts using terminal continuation. Lab Invest *84*, 131-137.

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A., 3rd, Kouzarides, T., Martignetti, J. A., and et al. (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol *154*, 125-169.

Chen, Y., Kamat, V., Dougherty, E. R., Bittner, M. L., Meltzer, P. S., and Trent, J. M. (2002). Ratio statistics of gene expression levels and applications to microarray data analysis. Bioinformatics *18*, 1207-1215.

Chicurel, M. (2002). Bioinformatics: bringing it all together. Nature *419*, 751, 753, 755 passim.

Chuck, A. S., Clarke, M. F., and Palsson, B. O. (1996). Retroviral infection is limited by Brownian motion. Hum Gene Ther *7*, 1527-1534.

Churchill, G. A. (2004). Using ANOVA to analyze microarray data. Biotechniques *37*, 173-175, 177.

Clark, D. A., Freeland, M. L., Mackie, L. K., Jarrett, R. F., and Onions, D. E. (1993). Prevalence of antibody to human herpesvirus 7 by age. J Infect Dis *168*, 251-252.

Compton, T., Nowlin, D. M., and Cooper, N. R. (1993). Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. Virology *193*, 834-841.

Corbeil, J., Sheeter, D., Genini, D., Rought, S., Leoni, L., Du, P., Ferguson, M., Masys, D. R., Welsh, J. B., Fink, J. L., *et al.* (2001). Temporal gene regulation during HIV-1 infection of human CD4+ T cells. Genome Res *11*, 1198-1204.

Cossman, J. (2001). Gene expression analysis of single neoplastic cells and the pathogenesis of Hodgkin's lymphoma. J Histochem Cytochem *49*, 799-800.

Cotter, P. A., and Miller, J. F. (1998). In vivo and ex vivo regulation of bacterial virulence gene expression. Curr Opin Microbiol *1*, 17-26.

Covell, D. G., Wallqvist, A., Rabow, A. A., and Thanki, N. (2003). Molecular classification of cancer: unsupervised self-organizing map analysis of gene expression microarray data. Mol Cancer Ther *2*, 317-332.

Crowley, R. W., Secchiero, P., Zella, D., Cara, A., Gallo, R. C., and Lusso, P. (1996). Interference between human herpesvirus 7 and HIV-1 in mononuclear phagocytes. J Immunol *156*, 2004-2008.

Cuadras, M. A., Feigelstock, D. A., An, S., and Greenberg, H. B. (2002). Gene expression pattern in Caco-2 cells following rotavirus infection. J Virol *76*, 4467-4482.

Cummings, C. A., and Relman, D. A. (2000). Using DNA microarrays to study host-microbe interactions. Emerg Infect Dis *6*, 513-525.

Cuomo, L., Trivedi, P., de Grazia, U., Calogero, A., D'Onofrio, M., Yang, W., Frati, L., Faggioni, A., Rymo, L., and Ragona, G. (1998). Upregulation of Epstein-Barr virus-encoded latent membrane protein by human herpesvirus 6 superinfection of EBV-carrying Burkitt lymphoma cells. J Med Virol *55*, 219-226.

Dana, A. N., Hong, Y. S., Kern, M. K., Hillenmeyer, M. E., Harker, B. W., Lobo, N. F., Hogan, J. R., Romans, P., and Collins, F. H. (2005). Gene expression patterns associated with blood-feeding in the malaria mosquito Anopheles gambiae. BMC Genomics *6*, 5.

Das, S., Sierra, J. C., Soman, K. V., Suarez, G., Mohammad, A. A., Dang, T. A., Luxon, B. A., and Reyes, V. E. (2005). Differential protein expression profiles of gastric epithelial cells following Helicobacter pylori infection using ProteinChips. J Proteome Res *4*, 920-930.

de Brevern, A. G., Hazout, S., and Malpertuy, A. (2004). Influence of microarrays experiments missing values on the stability of gene groups by hierarchical clustering. BMC Bioinformatics *5*, 114.

de La Fuente, C., Deng, L., Santiago, F., Arce, L., Wang, L., and Kashanchi, F. (2000). Gene expression array of HTLV type 1-infected T cells: Up-regulation of transcription factors and cell cycle genes. AIDS Res Hum Retroviruses *16*, 1695-1700.

de la Fuente, C., Santiago, F., Deng, L., Eadie, C., Zilberman, I., Kehn, K., Maddukuri, A., Baylor, S., Wu, K., Lee, C. G., *et al.* (2002). Gene expression profile of HIV-1 Tat expressing cells: a close interplay between proliferative and differentiation signals. BMC Biochem *3*, 14.

de Saizieu, A., Gardes, C., Flint, N., Wagner, C., Kamber, M., Mitchell, T. J., Keck, W., Amrein, K. E., and Lange, R. (2000). Microarray-based identification of a novel Streptococcus pneumoniae regulon controlled by an autoinduced peptide. J Bacteriol *182*, 4696-4703.

DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. (1996). Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet *14*, 457-460.

DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science *278*, 680-686.

Diehn, M., and Relman, D. A. (2001). Comparing functional genomic datasets: lessons from DNA microarray analyses of host-pathogen interactions. Curr Opin Microbiol *4*, 95-101.

Domachowske, J. B., Bonville, C. A., Easton, A. J., and Rosenberg, H. F. (2002). Differential expression of proinflammatory cytokine genes in vivo in response to pathogenic and nonpathogenic pneumovirus infections. J Infect Dis *186*, 8-14.

Drago, F., Ranieri, E., Malaguti, F., Losi, E., and Rebora, A. (1997). Human herpesvirus 7 in pityriasis rosea. Lancet *349*, 1367-1368.

Easton, A. J., and Clements, J. B. (1980). Temporal regulation of herpes simplex virus type 2 transcription and characterization of virus immediate early mRNA's. Nucleic Acids Res *8*, 2627-2645.

Ebrahimi, B., Dutia, B. M., Roberts, K. L., Garcia-Ramirez, J. J., Dickinson, P., Stewart, J. P., Ghazal, P., Roy, D. J., and Nash, A. A. (2003). Transcriptome profile of murine gammaherpesvirus-68 lytic infection. J Gen Virol *84*, 99-109.

Eckmann, L., Smith, J. R., Housley, M. P., Dwinell, M. B., and Kagnoff, M. F. (2000). Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria Salmonella. J Biol Chem *275*, 14084-14094.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Fakhari, F. D., and Dittmer, D. P. (2002). Charting latency transcripts in Kaposi's sarcoma-associated herpesvirus by whole-genome real-time quantitative PCR. J Virol *76*, 6213-6223.

Fernandez, C., Boutolleau, D., Manichanh, C., Mangeney, N., Agut, H., and Gautheret-Dejean, A. (2002). Quantitation of HHV-7 genome by real-time polymerase chain reaction assay using MGB probe technology. J Virol Methods *106*, 11-16.

Fielden, M. R., Halgren, R. G., Dere, E., and Zacharewski, T. R. (2002). GP3: GenePix post-processing program for automated analysis of raw microarray data. Bioinformatics *18*, 771-773.

Florens, L., Washburn, M. P., Raine, J. D., Anthony, R. M., Grainger, M., Haynes, J. D., Moch, J. K., Muster, N., Sacci, J. B., Tabb, D. L., *et al.* (2002). A proteomic view of the Plasmodium falciparum life cycle. Nature *419*, 520-526.

Frenkel, N., Schirmer, E. C., Wyatt, L. S., Katsafanas, G., Roffman, E., Danovich, R. M., and June, C. H. (1990). Isolation of a new herpesvirus from human CD4+ T cells. Proc Natl Acad Sci U S A *87*, 748-752.

Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci U S A *85*, 8998-9002.

Fujisaki, H., Tanaka-Taya, K., Tanabe, H., Hara, T., Miyoshi, H., Okada, S., and Yamanishi, K. (1998). Detection of human herpesvirus 7 (HHV-7) DNA in breast milk by polymerase chain reaction and prevalence of HHV-7 antibody in breast-fed and bottle-fed children. J Med Virol *56*, 275-279.

Furukawa, M., Yasukawa, M., Yakushijin, Y., and Fujita, S. (1994). Distinct effects of human herpesvirus 6 and human herpesvirus 7 on surface molecule expression and function of CD4+ T cells. J Immunol *152*, 5768-5775.

Geballe, A. P., Leach, F. S., and Mocarski, E. S. (1986). Regulation of cytomegalovirus late gene expression: gamma genes are controlled by posttranscriptional events. J Virol *57*, 864-874.

Geiss, G. K., An, M. C., Bumgarner, R. E., Hammersmark, E., Cunningham, D., and Katze, M. G. (2001). Global impact of influenza virus on cellular pathways is mediated by both replication-dependent and -independent events. J Virol *75*, 4321-4331.

Geiss, G. K., Bumgarner, R. E., An, M. C., Agy, M. B., van 't Wout, A. B., Hammersmark, E., Carter, V. S., Upchurch, D., Mullins, J. I., and Katze, M. G. (2000). Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. Virology *266*, 8-16.

Geiss, G. K., Carter, V. S., He, Y., Kwieciszewski, B. K., Holzman, T., Korth, M. J., Lazaro, C. A., Fausto, N., Bumgarner, R. E., and Katze, M. G. (2003). Gene expression profiling of the cellular transcriptional network regulated by alpha/beta interferon and its partial attenuation by the hepatitis C virus nonstructural 5A protein. J Virol *77*, 6367-6375.

Geng, Y. Q., Chandran, B., Josephs, S. F., and Wood, C. (1992). Identification and characterization of a human herpesvirus 6 gene segment that trans activates the human immunodeficiency virus type 1 promoter. J Virol *66*, 1564-1570.

Gerhold, D., Lu, M., Xu, J., Austin, C., Caskey, C. T., and Rushmore, T. (2001). Monitoring expression of genes involved in drug metabolism and toxicology using DNA microarrays. Physiol Genomics *5*, 161-170.

Gibson, W. (1983). Protein counterparts of human and simian cytomegaloviruses. Virology *128*, 391-406.

Goebel, S. J., Johnson, G. P., Perkus, M. E., Davis, S. W., Winslow, J. P., and Paoletti, E. (1990). The complete DNA sequence of vaccinia virus. Virology *179*, 247-266, 517-263.

Goff, L. A., Bowers, J., Schwalm, J., Howerton, K., Getts, R. C., and Hart, R. P. (2004). Evaluation of sense-strand mRNA amplification by comparative quantitative PCR. BMC Genomics *5*, 76.

Gompels, U. A., Nicholas, J., Lawrence, G., Jones, M., Thomson, B. J., Martin, M. E., Efstathiou, S., Craxton, M., and Macaulay, H. A. (1995). The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. Virology *209*, 29-51.

Graveel, C. R., Jatkoe, T., Madore, S. J., Holt, A. L., and Farnham, P. J. (2001). Expression profiling and identification of novel genes in hepatocellular carcinomas. Oncogene *20*, 2704-2712.

Gray, N. S., Wodicka, L., Thunnissen, A. M., Norman, T. C., Kwon, S., Espinoza, F. H., Morgan, D. O., Barnes, G., LeClerc, S., Meijer, L., *et al.* (1998). Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. Science *281*, 533-538.

Guerra, S., Lopez-Fernandez, L. A., Conde, R., Pascual-Montano, A., Harshman, K., and Esteban, M. (2004). Microarray analysis reveals characteristic changes of host cell gene expression in response to attenuated modified vaccinia virus Ankara infection of human HeLa cells. J Virol *78*, 5820-5834.

Guerra, S., Lopez-Fernandez, L. A., Pascual-Montano, A., Munoz, M., Harshman, K., and Esteban, M. (2003). Cellular gene expression survey of vaccinia virus infection of human HeLa cells. J Virol *77*, 6493-6506.

Hacia, J. G. (1999). Resequencing and mutational analysis using oligonucleotide microarrays. Nat Genet *21*, 42-47.

Hakenbeck, R., Balmelle, N., Weber, B., Gardes, C., Keck, W., and de Saizieu, A. (2001). Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of Streptococcus pneumoniae. Infect Immun *69*, 2477-2486.

Hannon, G. J. (2002). RNA interference. Nature *418*, 244-251.

Harrington, C. A., Rosenow, C., and Retief, J. (2000). Monitoring gene expression using DNA microarrays. Curr Opin Microbiol *3*, 285-291.

Harris, T. M., and Childs, G. (2002). Global gene expression patterns during differentiation of F9 embryonal carcinoma cells into parietal endoderm. Funct Integr Genomics *2*, 105-119.

Hauser, M. A., Li, Y. J., Xu, H., Noureddine, M. A., Shao, Y. S., Gullans, S. R., Scherzer, C. R., Jensen, R. V., McLaurin, A. C., Gibson, J. R., *et al.* (2005). Expression profiling of substantia nigra in

Parkinson disease, progressive supranuclear palsy, and frontotemporal dementia with parkinsonism. Arch Neurol *62*, 917-921.

Hayashi, M. L., Blankenship, C., and Shenk, T. (2000). Human cytomegalovirus UL69 protein is required for efficient accumulation of infected cells in the G1 phase of the cell cycle. Proc Natl Acad Sci U S A *97*, 2692-2696.

Hayward, R. E. (2000). Plasmodium falciparum phosphoenolpyruvate carboxykinase is developmentally regulated in gametocytes. Mol Biochem Parasitol *107*, 227-240.

Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P., *et al.* (2001). Gene-expression profiles in hereditary breast cancer. N Engl J Med *344*, 539-548.

Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., Hughes, J. E., Snesrud, E., Lee, N., and Quackenbush, J. (2000). A concise guide to cDNA microarray analysis. Biotechniques *29*, 548-550, 552-544, 556 passim.

Heller, M. J. (2002). DNA microarray technology: devices, systems, and applications. Annu Rev Biomed Eng *4*, 129-153.

Helmberg, A. (2001). DNA-microarrays: novel techniques to study aging and guide gerontologic medicine. Exp Gerontol *36*, 1189-1198.

Hertel, L., and Mocarski, E. S. (2004). Global Analysis of Host Cell Gene Expression Late during Cytomegalovirus Infection Reveals Extensive Dysregulation of Cell Cycle Gene Expression and Induction of Pseudomitosis Independent of US28 Function. J Virol *78*, 11988-12011.

Hinchliffe, S. J., Isherwood, K. E., Stabler, R. A., Prentice, M. B., Rakin, A., Nichols, R. A., Oyston, P. C., Hinds, J., Titball, R. W., and Wren, B. W. (2003). Application of DNA microarrays to study the evolutionary genomics of Yersinia pestis and Yersinia pseudotuberculosis. Genome Res *13*, 2018-2029.

Ho, W. Z., Cherukuri, R., Ge, S. D., Cutilli, J. R., Song, L., Whitko, S., and Douglas, S. D. (1993). Centrifugal enhancement of human immunodeficiency virus type 1 infection and human cytomegalovirus gene expression in human primary monocyte/macrophages in vitro. J Leukoc Biol *53*, 208-212.

Hobbs, W. E., 2nd, and DeLuca, N. A. (1999). Perturbation of cell cycle progression and cellular gene expression as a function of herpes simplex virus ICP0. J Virol *73*, 8245-8255.

Hodgkin, P. D., Scalzo, A. A., Swaminathan, N., Price, P., and Shellam, G. R. (1988). Murine cytomegalovirus binds reversibly to mouse embryo fibroblasts: implications for quantitation and explanation of centrifugal enhancement. J Virol Methods *22*, 215-230.

Honda, M., Kaneko, S., Kawai, H., Shirota, Y., and Kobayashi, K. (2001). Differential gene expression between chronic hepatitis B and C hepatic lesion. Gastroenterology *120*, 955-966.

Honda, M., Kawai, H., Shirota, Y., Yamashita, T., and Kaneko, S. (2005). Differential gene expression profiles in stage I primary biliary cirrhosis. Am J Gastroenterol *100*, 2019-2030.

Honess, R. W., and Roizman, B. (1974). Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J Virol *14*, 8-19.

Honess, R. W., and Watson, D. H. (1977). Herpes simplex virus resistance and sensitivity to phosphonoacetic acid. J Virol *21*, 584-600.

Huang, S., and Hanson, L. A. (1998). Temporal gene regulation of the channel catfish virus (Ictalurid herpesvirus 1). J Virol *72*, 1910-1917.

Hudson, J. B. (1988). Further studies on the mechanism of centrifugal enhancement of cytomegalovirus infectivity. J Virol Methods *19*, 97-108.

Hudson, J. B., Misra, V., and Mosmann, T. R. (1976). Cytomegalovirus infectivity: analysis of the phenomenon of centrifugal enhancement of infectivity. Virology *72*, 235-243.

Hummerich, L., Muller, R., Hess, J., Kokocinski, F., Hahn, M., Furstenberger, G., Mauch, C., Lichter, P., and Angel, P. (2005). Identification of novel tumour-associated genes differentially expressed in the process of squamous cell cancer development. Oncogene.

Iizuka, N., Oka, M., Yamada-Okabe, H., Mori, N., Tamesa, T., Okada, T., Takemoto, N., Sakamoto, K., Hamada, K., Ishitsuka, H., *et al.* (2005). Self-organizing-map-based molecular signature representing the development of hepatocellular carcinoma. FEBS Lett *579*, 1089-1100.

Imanishi, T., Itoh, T., Suzuki, Y., O'Donovan, C., Fukuchi, S., Koyanagi, K. O., Barrero, R. A., Tamura, T., Yamaguchi-Kabata, Y., Tanino, M., *et al.* (2004). Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biol *2*, e162.

Inoue, N., Winter, J., Lal, R. B., Offermann, M. K., and Koyano, S. (2003). Characterization of entry mechanisms of human herpesvirus 8 by using an Rta-dependent reporter cell line. J Virol *77*, 8147-8152.

Irizarry, R. A., Warren, D., Spencer, F., Kim, I. F., Biswal, S., Frank, B. C., Gabrielson, E., Garcia, J. G., Geoghegan, J., Germino, G., *et al.* (2005). Multiple-laboratory comparison of microarray platforms. Nat Methods *2*, 345-350.

Iskenderian, A. C., Huang, L., Reilly, A., Stenberg, R. M., and Anders, D. G. (1996). Four of eleven loci required for transient complementation of human cytomegalovirus DNA replication cooperate to activate expression of replication genes. J Virol *70*, 383-392.

Itoh, K., Kawasaki, S., Kawamoto, S., Seishima, M., Chiba, H., Michibata, H., Wakimoto, K., Imai, Y., Minesaki, Y., Otsuji, M., and Okubo, K. (2005). Identification of differentially expressed genes in psoriasis using expression profiling approaches. Exp Dermatol *14*, 667-674.

Jenner, R. G., Alba, M. M., Boshoff, C., and Kellam, P. (2001). Kaposi's sarcoma-associated herpesvirus latent and lytic gene expression as revealed by DNA arrays. J Virol *75*, 891-902.

Jenner, R. G., and Young, R. A. (2005). Insights into host responses against pathogens from transcriptional profiling. Nat Rev Microbiol *3*, 281-294.

Johannes, G., Carter, M. S., Eisen, M. B., Brown, P. O., and Sarnow, P. (1999). Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proc Natl Acad Sci U S A *96*, 13118-13123.

Jones, J. O., and Arvin, A. M. (2003). Microarray analysis of host cell gene transcription in response to varicella-zoster virus infection of human T cells and fibroblasts in vitro and SCIDhu skin xenografts in vivo. J Virol *77*, 1268-1280.

Kane, M. D., Jatkoe, T. A., Stumpf, C. R., Lu, J., Thomas, J. D., and Madore, S. J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res *28*, 4552-4557.

Kato-Maeda, M., Rhee, J. T., Gingeras, T. R., Salamon, H., Drenkow, J., Smittipat, N., and Small, P. M. (2001). Comparing genomes within the species Mycobacterium tuberculosis. Genome Res *11*, 547-554.

Keil, G. M., Ebeling-Keil, A., and Koszinowski, U. H. (1984). Temporal regulation of murine cytomegalovirus transcription and mapping of viral RNA synthesized at immediate early times after infection. J Virol *50*, 784-795.

Kellam, P. (2001). Post-genomic virology: the impact of bioinformatics, microarrays and proteomics on investigating host and pathogen interactions. Rev Med Virol *11*, 313-329.

Kempf, W., Adams, V., Kleinhans, M., Burg, G., Panizzon, R. G., Campadelli-Fiume, G., and Nestle, F. O. (1999). Pityriasis rosea is not associated with human herpesvirus 7. Arch Dermatol *135*, 1070-1072.

Kempf, W., Adams, V., Mirandola, P., Menotti, L., Di Luca, D., Wey, N., Muller, B., and Campadelli-Fiume, G. (1998). Persistence of human herpesvirus 7 in normal tissues detected by expression of a structural antigen. J Infect Dis *178*, 841-845.

Kennedy, P. G., Grinfeld, E., Craigon, M., Vierlinger, K., Roy, D., Forster, T., and Ghazal, P. (2005). Transcriptomal analysis of varicella-zoster virus infection using long oligonucleotide-based microarrays. J Gen Virol *86*, 2673-2684.

Kerr, M. K., Martin, M., and Churchill, G. A. (2000). Analysis of variance for gene expression microarray data. J Comput Biol *7*, 819-837.

Khodarev, N. N., Advani, S. J., Gupta, N., Roizman, B., and Weichselbaum, R. R. (1999). Accumulation of specific RNAs encoding transcriptional factors and stress response proteins against a background of severe depletion of cellular RNAs in cells infected with herpes simplex virus 1. Proc Natl Acad Sci U S A *96*, 12062-12067.

Kitahara, O., Furukawa, Y., Tanaka, T., Kihara, C., Ono, K., Yanagawa, R., Nita, M. E., Takagi, T., Nakamura, Y., and Tsunoda, T. (2001). Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. Cancer Res *61*, 3544-3549.

Kotani, H., Newton, P. B., 3rd, Zhang, S., Chiang, Y. L., Otto, E., Weaver, L., Blaese, R. M., Anderson, W. F., and McGarrity, G. J. (1994). Improved methods of retroviral vector transduction and production for gene therapy. Hum Gene Ther *5*, 19-28.

Kothapalli, R., Yoder, S. J., Mane, S., and Loughran, T. P., Jr. (2002). Microarray results: how accurate are they? BMC Bioinformatics *3*, 22.

Kuwabara, P. E., and Coulson, A. (2000). RNAi--prospects for a general technique for determining gene function. Parasitol Today *16*, 347-349.

LaBaer, J., and Ramachandran, N. (2005). Protein microarrays as tools for functional proteomics. Curr Opin Chem Biol *9*, 14-19.

Landazuri, N., and Le Doux, J. M. (2004). Complexation of retroviruses with charged polymers enhances gene transfer by increasing the rate that viruses are delivered to cells. J Gene Med *6*, 1304-1319.

Larkin, J. E., Frank, B. C., Gavras, H., Sultana, R., and Quackenbush, J. (2005). Independence and reproducibility across microarray platforms. Nat Methods *2*, 337-344.

Larsson, O., Wahlestedt, C., and Timmons, J. A. (2005). Considerations when using the significance analysis of microarrays (SAM) algorithm. BMC Bioinformatics *6*, 129.

Leong, W. F., Tan, H. C., Ooi, E. E., Koh, D. R., and Chow, V. T. (2005). Microarray and real-time RT-PCR analyses of differential human gene expression patterns induced by severe acute respiratory syndrome (SARS) coronavirus infection of Vero cells. Microbes Infect *7*, 248-259.

Levesque, K., Finzi, A., Binette, J., and Cohen, E. A. (2004). Role of CD4 receptor down-regulation during HIV-1 infection. Curr HIV Res *2*, 51-59.

Lieberman, J., Song, E., Lee, S. K., and Shankar, P. (2003). Interfering with disease: opportunities and roadblocks to harnessing RNA interference. Trends Mol Med *9*, 397-403.

Little, S. J., Riggs, N. L., Chowers, M. Y., Fitch, N. J., Richman, D. D., Spina, C. A., and Guatelli, J. C. (1994). Cell surface CD4 downregulation and resistance to superinfection induced by a defective provirus of HIV-1. Virology *205*, 578-582.

Liu, B., and Stinski, M. F. (1992). Human cytomegalovirus contains a tegument protein that enhances transcription from promoters with upstream ATF and AP-1 cis-acting elements. J Virol *66*, 4434-4444.

Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E. L. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol *14*, 1675-1680.

Lockhart, D. J., and Winzeler, E. A. (2000). Genomics, gene expression and DNA arrays. Nature *405*, 827-836.

Lopetegui, P., Campo-Vera, H., and Yamanishi, K. (1985). Varicella-zoster virus (VZV)-specific polypeptides detected in cells treated with metabolic inhibitors. Microbiol Immunol *29*, 569-575.

Lundquist, C. A., Tobiume, M., Zhou, J., Unutmaz, D., and Aiken, C. (2002). Nef-mediated downregulation of CD4 enhances human immunodeficiency virus type 1 replication in primary T lymphocytes. J Virol *76*, 4625-4633.

Luo, L., Salunga, R. C., Guo, H., Bittner, A., Joy, K. C., Galindo, J. E., Xiao, H., Rogers, K. E., Wan, J. S., Jackson, M. R., and Erlander, M. G. (1999). Gene expression profiles of laser-captured adjacent neuronal subtypes. Nat Med *5*, 117-122.

Lusso, P., Secchiero, P., Crowley, R. W., Garzino-Demo, A., Berneman, Z. N., and Gallo, R. C. (1994). CD4 is a critical component of the receptor for human herpesvirus 7: interference with human immunodeficiency virus. Proc Natl Acad Sci U S A *91*, 3872-3876.

MacFarlane, R., Bhattacharya, D., and Singh, U. (2005a). Genomic DNA microarrays for Entamoeba histolytica: applications for use in expression profiling and strain genotyping. Exp Parasitol *110*, 196-202.

MacFarlane, R. C., Shah, P. H., and Singh, U. (2005b). Transcriptional profiling of Entamoeba histolytica trophozoites. Int J Parasitol *35*, 533-542.

Marko, N. F., Frank, B., Quackenbush, J., and Lee, N. H. (2005). A robust method for the amplification of RNA in the sense orientation. BMC Genomics *6*, 27.

Mayne, M., Cheadle, C., Soldan, S. S., Cermelli, C., Yamano, Y., Akhyani, N., Nagel, J. E., Taub, D. D., Becker, K. G., and Jacobson, S. (2001). Gene expression profile of herpesvirus-infected T cells obtained using immunomicroarrays: induction of proinflammatory mechanisms. J Virol *75*, 11641-11650.

McClintick, J. N., Jerome, R. E., Nicholson, C. R., Crabb, D. W., and Edenberg, H. J. (2003). Reproducibility of oligonucleotide arrays using small samples. BMC Genomics *4*, 4.

McPherson, J. D., Marra, M., Hillier, L., Waterston, R. H., Chinwalla, A., Wallis, J., Sekhon, M., Wylie, K., Mardis, E. R., Wilson, R. K., *et al.* (2001). A physical map of the human genome. Nature *409*, 934-941.

Megaw, A. G., Rapaport, D., Avidor, B., Frenkel, N., and Davison, A. J. (1998). The DNA sequence of the RK strain of human herpesvirus 7. Virology *244*, 119-132.

Mendez, J. C., Dockrell, D. H., Espy, M. J., Smith, T. F., Wilson, J. A., Harmsen, W. S., Ilstrup, D., and Paya, C. V. (2001). Human beta-herpesvirus interactions in solid organ transplant recipients. J Infect Dis *183*, 179-184.

Menegazzi, P., Galvan, M., Rotola, A., Ravaioli, T., Gonelli, A., Cassai, E., and Di Luca, D. (1999). Temporal mapping of transcripts in human herpesvirus-7. J Gen Virol *80 (Pt 10)*, 2705-2712.

Merrell, D. S., Butler, S. M., Qadri, F., Dolganov, N. A., Alam, A., Cohen, M. B., Calderwood, S. B., Schoolnik, G. K., and Camilli, A. (2002). Host-induced epidemic spread of the cholera bacterium. Nature *417*, 642-645.

Miller, J. C., Zhou, H., Kwekel, J., Cavallo, R., Burke, J., Butler, E. B., Teh, B. S., and Haab, B. B. (2003). Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers. Proteomics *3*, 56-63.

Mirandola, P., Menegazzi, P., Merighi, S., Ravaioli, T., Cassai, E., and Di Luca, D. (1998). Temporal mapping of transcripts in herpesvirus 6 variants. J Virol *72*, 3837-3844.

Mitchell, R., Chiang, C. Y., Berry, C., and Bushman, F. (2003). Global analysis of cellular transcription following infection with an HIV-based vector. Mol Ther *8*, 674-687.

Mocarski, E. S., Courcelle, C.T (2001). Cytomegaloviruses and their replication. In Fields Virology, D. M. Knipe, Howley, P. M., ed. (Philadelphia, Lippincott Williams & Wilkins), pp. 2629-2674.

Morgan, F. W., Stewart, J. A., Smith, A. N., and Tarnuzzer, R. W. (2005). Differential expression of stress response genes in the H-Tx rat model of congenital hydrocephalus. Brain Res Mol Brain Res *138*, 273-290.

Morgan, R. W., Sofer, L., Anderson, A. S., Bernberg, E. L., Cui, J., and Burnside, J. (2001). Induction of host gene expression following infection of chicken embryo fibroblasts with oncogenic Marek's disease virus. J Virol *75*, 533-539.

Moses, A. V., Jarvis, M. A., Raggo, C., Bell, Y. C., Ruhl, R., Luukkonen, B. G., Griffith, D. J., Wait, C. L., Druker, B. J., Heinrich, M. C., *et al.* (2002). A functional genomics approach to Kaposi's sarcoma. Ann N Y Acad Sci *975*, 180-191.

Mossman, K. L., Macgregor, P. F., Rozmus, J. J., Goryachev, A. B., Edwards, A. M., and Smiley, J. R. (2001). Herpes simplex virus triggers and then disarms a host antiviral response. J Virol *75*, 750-758.

Murphy, N., Millar, E., and Lee, C. S. (2005). Gene expression profiling in breast cancer: towards individualising patient management. Pathology *37*, 271-277.

Nam, D. K., Lee, S., Zhou, G., Cao, X., Wang, C., Clark, T., Chen, J., Rowley, J. D., and Wang, S. M. (2002). Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. Proc Natl Acad Sci U S A *99*, 6152-6156.

Nicholas, J. (1996). Determination and analysis of the complete nucleotide sequence of human herpesvirus 7. J Virol *70*, 5975-5989.

Nyberg-Hoffman, C., Shabram, P., Li, W., Giroux, D., and Aguilar-Cordova, E. (1997). Sensitivity and reproducibility in adenoviral infectious titer determination. Nat Med *3*, 808-811.

O'Doherty, U., Swiggard, W. J., and Malim, M. H. (2000). Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. J Virol *74*, 10074-10080.

Ohyashiki, J. H., Takaku, T., Ojima, T., Abe, K., Yamamoto, K., Zhang, Y., and Ohyashiki, K. (2005). Transcriptional profiling of human herpesvirus type B (HHV-6B) in an adult T cell leukemia cell line as in vitro model for persistent infection. Biochem Biophys Res Commun *329*, 11-17.

Okabe, H., Satoh, S., Kato, T., Kitahara, O., Yanagawa, R., Yamaoka, Y., Tsunoda, T., Furukawa, Y., and Nakamura, Y. (2001). Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res *61*, 2129-2137.

Okada, Y., Sahara, T., Mitsubayashi, H., Ohgiya, S., and Nagashima, T. (2005). Knowledge-assisted recognition of cluster boundaries in gene expression data. Artif Intell Med *35*, 171-183.

Okamoto, T., Suzuki, T., and Yamamoto, N. (2000). Microarray fabrication with covalent attachment of DNA using bubble jet technology. Nat Biotechnol *18*, 438-441.

Ordway, J. M., Fenster, S. D., Ruan, H., and Curran, T. (2005). A transcriptome map of cellular transformation by the fos oncogene. Mol Cancer *4*, 19.

Osborn, J. E., and Walker, D. L. (1968). Enhancement of infectivity of murine cytomegalovirus in vitro by centrifugal inoculation. J Virol *2*, 853-858.

Osman, H. K., Peiris, J. S., Taylor, C. E., Warwicker, P., Jarrett, R. F., and Madeley, C. R. (1996). "Cytomegalovirus disease" in renal allograft recipients: is human herpesvirus 7 a co-factor for disease progression? J Med Virol *48*, 295-301.

Pan, L. Z., Werner, A., and Levy, J. A. (1993). Detection of plasma viremia in human immunodeficiency virus-infected individuals at all clinical stages. J Clin Microbiol *31*, 283-288.

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. Nature *405*, 837-846.

Pandya, U., Allen, C. A., Watson, D. A., and Niesel, D. W. (2005). Global profiling of Streptococcus pneumoniae gene expression at different growth temperatures. Gene *360*, 45-54.

Pang, A. L., Johnson, W., Ravindranath, N., Dym, M., Rennert, O. M., and Chan, W. Y. (2005). Expression Profiling of Purified Male Germ Cells: Stage Specific Expression Patterns Related to Meiosis and Post-Meiotic Development. Physiol Genomics.

Pari, G. S., Kacica, M. A., and Anders, D. G. (1993). Open reading frames UL44, IRS1/TRS1, and UL36-38 are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA synthesis. J Virol *67*, 2575-2582.

Paulose-Murphy, M., Ha, N. K., Xiang, C., Chen, Y., Gillim, L., Yarchoan, R., Meltzer, P., Bittner, M., Trent, J., and Zeichner, S. (2001). Transcription program of human herpesvirus 8 (kaposi's sarcoma-associated herpesvirus). J Virol *75*, 4843-4853.

Pease, A. C., Solas, D., Sullivan, E. J., Cronin, M. T., Holmes, C. P., and Fodor, S. P. (1994). Light-generated oligonucleotide arrays for rapid DNA sequence analysis. Proc Natl Acad Sci U S A *91*, 5022-5026.

Peden, K. W. C., Martin, M. A. (1995). Virological and molecular genetic techniques for studies of established HIV isolates. In HIV: A Practical Approach - Virology & Immunology, J. Karn, ed. (New York, Oxford University Press), pp. 21-46.

Pellett, P. E., Dominguez, G. (2001). Human herpesviruses 6A, 6A, and 7 and their replication. In Fields Virology, D. M. Knipe, Howley, P. M., ed. (Philadelphia, Lippincott Williams & Wilkins), pp. 2769-2784.

Penfold, M. E., and Mocarski, E. S. (1997). Formation of cytomegalovirus DNA replication compartments defined by localization of viral proteins and DNA synthesis. Virology *239*, 46-61.

Petalidis, L., Bhattacharyya, S., Morris, G. A., Collins, V. P., Freeman, T. C., and Lyons, P. A. (2003). Global amplification of mRNA by template-switching PCR: linearity and application to microarray analysis. Nucleic Acids Res *31*, e142.

Petrik, J. (2001). Microarray technology: the future of blood testing? Vox Sang *80*, 1-11.

Pietiainen, V., Huttunen, P., and Hyypia, T. (2000). Effects of echovirus 1 infection on cellular gene expression. Virology *276*, 243-250.

Pietroboni, G. R., Harnett, G. B., and Bucens, M. R. (1989). Centrifugal enhancement of human immunodeficiency virus (HIV) and human herpesvirus type 6 (HHV-6) infection in vitro. J Virol Methods *24*, 85-90.

Piguet, V., Schwartz, O., Le Gall, S., and Trono, D. (1999). The downregulation of CD4 and MHC-I by primate lentiviruses: a paradigm for the modulation of cell surface receptors. Immunol Rev *168*, 51-63.

Poggioli, G. J., DeBiasi, R. L., Bickel, R., Jotte, R., Spalding, A., Johnson, G. L., and Tyler, K. L. (2002). Reovirus-induced alterations in gene expression related to cell cycle regulation. J Virol *76*, 2585-2594.

Polson, A. G., Wang, D., DeRisi, J., and Ganem, D. (2002). Modulation of host gene expression by the constitutively active G protein-coupled receptor of Kaposi's sarcoma-associated herpesvirus. Cancer Res *62*, 4525-4530.

Poole, L. J., Yu, Y., Kim, P. S., Zheng, Q. Z., Pevsner, J., and Hayward, G. S. (2002). Altered patterns of cellular gene expression in dermal microvascular endothelial cells infected with Kaposi's sarcoma-associated herpesvirus. J Virol *76*, 3395-3420.

Puskas, L. G., Zvara, A., Hackler, L., Jr., and Van Hummelen, P. (2002). RNA amplification results in reproducible microarray data with slight ratio bias. Biotechniques *32*, 1330-1334, 1336, 1338, 1340.

Qian, G., Wood, C., and Chandran, B. (1993). Identification and characterization of glycoprotein gH of human herpesvirus-6. Virology *194*, 380-386.

Quackenbush, J. (2001). Computational analysis of microarray data. Nat Rev Genet *2*, 418-427.

Quackenbush, J. (2002). Microarray data normalization and transformation. Nat Genet *32 Suppl*, 496-501.

Rajeevan, M. S., Dimulescu, I. M., Vernon, S. D., Verma, M., and Unger, E. R. (2003). Global amplification of sense RNA: a novel method to replicate and archive mRNA for gene expression analysis. Genomics *82*, 491-497.

Rapp, J. C., Krug, L. T., Inoue, N., Dambaugh, T. R., and Pellett, P. E. (2000). U94, the human herpesvirus 6 homolog of the parvovirus nonstructural gene, is highly conserved among isolates and is expressed at low mRNA levels as a spliced transcript. Virology *268*, 504-516.

Rathod, P. K., Ganesan, K., Hayward, R. E., Bozdech, Z., and DeRisi, J. L. (2002). DNA microarrays for malaria. Trends Parasitol *18*, 39-45.

Reading, P. C., Moore, J. B., and Smith, G. L. (2003). Steroid hormone synthesis by vaccinia virus suppresses the inflammatory response to infection. J Exp Med *197*, 1269-1278.

Reading, S. A., Edwards, M. J., and Dimmock, N. J. (2001). Increasing the efficiency of virus infectivity assays: small inoculum volumes are as effective as centrifugal enhancement. J Virol Methods *98*, 167-169.

Renne, R., Barry, C., Dittmer, D., Compitello, N., Brown, P. O., and Ganem, D. (2001). Modulation of cellular and viral gene expression by the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus. J Virol *75*, 458-468.

Richmond, C. S., Glasner, J. D., Mau, R., Jin, H., and Blattner, F. R. (1999). Genome-wide expression profiling in Escherichia coli K-12. Nucleic Acids Res *27*, 3821-3835.

Roizman, B., Knipe, D. M. (2001a). Herpes simplex viruses and their replication. In Fields Virology, D. M. Knipe, Howley, P. M., ed. (Philadelphia, Lippincott Williams & Wilkins), pp. 2399-2460.

Roizman, B., Pellet, P.E. (2001b). The family herpesviridae: A brief introduction. In Fields Virology, D. M. Knipe, Howley, P. M., ed. (Philadelphia, Lippincott Willams & Wilkins), pp. 2381-2398.

Romanowski, M. J., and Shenk, T. (1997). Characterization of the human cytomegalovirus irs1 and trs1 genes: a second immediate-early transcription unit within irs1 whose product antagonizes transcriptional activation. J Virol *71*, 1485-1496.

Rosenberger, C. M., Scott, M. G., Gold, M. R., Hancock, R. E., and Finlay, B. B. (2000). Salmonella typhimurium infection and lipopolysaccharide stimulation induce similar changes in macrophage gene expression. J Immunol *164*, 5894-5904.

Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., *et al.* (2000). Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet *24*, 227-235.

Rozen, S., Skaletsky, H. J. (2000). Primer3 on the WWW for general users and for biologist programmers. In Bioinformatics Methods and Protocols: Methods in Molecular Biology, S. Krawetz, Misener, S., ed. (Totowa, NJ, Humana Press), pp. 365-386.

Ryo, A., Suzuki, Y., Arai, M., Kondoh, N., Wakatsuki, T., Hada, A., Shuda, M., Tanaka, K., Sato, C., Yamamoto, M., and Yamamoto, N. (2000). Identification and characterization of differentially expressed mRNAs in HIV type 1-infected human T cells. AIDS Res Hum Retroviruses *16*, 995-1005.

Sada, E., Yasukawa, M., Ito, C., Takeda, A., Shiosaka, T., Tanioka, H., and Fujita, S. (1996). Detection of human herpesvirus 6 and human herpesvirus 7 in the submandibular gland, parotid gland, and lip salivary gland by PCR. J Clin Microbiol *34*, 2320-2321.

Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., *et al.* (2003). TM4: a free, open-source system for microarray data management and analysis. Biotechniques *34*, 374-378.

Salahuddin, S. Z., Ablashi, D. V., Markham, P. D., Josephs, S. F., Sturzenegger, S., Kaplan, M., Halligan, G., Biberfeld, P., Wong-Staal, F., Kramarsky, B., and et al. (1986). Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. Science *234*, 596-601.

Salama, N., Guillemin, K., McDaniel, T. K., Sherlock, G., Tompkins, L., and Falkow, S. (2000). A whole-genome microarray reveals genetic diversity among Helicobacter pylori strains. Proc Natl Acad Sci U S A *97*, 14668-14673.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A *74*, 5463-5467.

Sarisky, R. T., and Hayward, G. S. (1996). Evidence that the UL84 gene product of human cytomegalovirus is essential for promoting oriLyt-dependent DNA replication and formation of replication compartments in cotransfection assays. J Virol *70*, 7398-7413.

Sasada, A., Takaori-Kondo, A., Shirakawa, K., Kobayashi, M., Abudu, A., Hishizawa, M., Imada, K., Tanaka, Y., and Uchiyama, T. (2005). APOBEC3G targets human T-cell leukemia virus type 1. Retrovirology *2*, 32.

Satterthwaite, G., Francis, S. E., Suvarna, K., Blakemore, S., Ward, C., Wallace, D., Braddock, M., and Crossman, D. (2005). Differential gene expression in coronary arteries from patients presenting with ischemic heart disease: further evidence for the inflammatory basis of atherosclerosis. Am Heart J *150*, 488-499.

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467-470.

Scherf, U., Ross, D. T., Waltham, M., Smith, L. H., Lee, J. K., Tanabe, L., Kohn, K. W., Reinhold, W. C., Myers, T. G., Andrews, D. T., *et al.* (2000). A gene expression database for the molecular pharmacology of cancer. Nat Genet *24*, 236-244.

Schindler, H., Wiese, A., Auer, J., and Burtscher, H. (2005). cRNA target preparation for microarrays: comparison of gene expression profiles generated with different amplification procedures. Anal Biochem *344*, 92-101.

Seal, B. S., Irving, J. M., and Whetstone, C. A. (1991). Transcriptional analysis of the bovine herpesvirus 1 Cooper isolate. Temporal analysis and characterization of immediate-early, early, and late RNA. Arch Virol *121*, 55-73.

Secchiero, P., Berneman, Z. N., Gallo, R. C., and Lusso, P. (1994). Biological and molecular characteristics of human herpesvirus 7: in vitro growth optimization and development of a syncytia inhibition test. Virology *202*, 506-512.

Secchiero, P., Bertolaso, L., Casareto, L., Gibellini, D., Vitale, M., Bemis, K., Aleotti, A., Capitani, S., Franchini, G., Gallo, R. C., and Zauli, G. (1998a). Human herpesvirus 7 infection induces profound cell cycle perturbations coupled to disregulation of cdc2 and cyclin B and polyploidization of CD4(+) T cells. Blood *92*, 1685-1696.

Secchiero, P., Bertolaso, L., Gibellini, D., Ricci, D., Bemis, K., Capitani, S., Gallo, R. C., and Zauli, G. (1998b). Enforced expression of human bcl-2 in CD4+ T cells enhances human herpesvirus 7 replication and induction of cytopathic effects. Eur J Immunol *28*, 1587-1596.

Secchiero, P., Flamand, L., Gibellini, D., Falcieri, E., Robuffo, I., Capitani, S., Gallo, R. C., and Zauli, G. (1997a). Human Herpesvirus 7 induces CD4(+) T-cell death by two distinct mechanisms: necrotic lysis in productively infected cells and apoptosis in uninfected or nonproductively infected cells. Blood *90*, 4502-4512.

Secchiero, P., Gibellini, D., Flamand, L., Robuffo, I., Marchisio, M., Capitani, S., Gallo, R. C., and Zauli, G. (1997b). Human herpesvirus 7 induces the down-regulation of CD4 antigen in lymphoid T cells without affecting p56lck levels. J Immunol *159*, 3412-3423.

Secchiero, P., Mirandola, P., Zella, D., Celeghini, C., Gonelli, A., Vitale, M., Capitani, S., and Zauli, G. (2001). Human herpesvirus 7 induces the functional up-regulation of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) coupled to TRAIL-R1 down-modulation in CD4(+) T cells. Blood *98*, 2474-2481.

Secchiero, P., Sun, D., De Vico, A. L., Crowley, R. W., Reitz, M. S., Jr., Zauli, G., Lusso, P., and Gallo, R. C. (1997c). Role of the extracellular domain of human herpesvirus 7 glycoprotein B in virus binding to cell surface heparan sulfate proteoglycans. J Virol *71*, 4571-4580.

Secchiero, P., Zella, D., Barabitskaja, O., Reitz, M. S., Capitani, S., Gallo, R. C., and Zauli, G. (1998c). Progressive and persistent downregulation of surface CXCR4 in CD4(+) T cells infected with human herpesvirus 7. Blood *92*, 4521-4528.

Seth, D., Gorrell, M. D., McGuinness, P. H., Leo, M. A., Lieber, C. S., McCaughan, G. W., and Haber, P. S. (2003). SMART amplification maintains representation of relative gene expression: quantitative validation by real time PCR and application to studies of alcoholic liver disease in primates. J Biochem Biophys Methods *55*, 53-66.

Shafti-Keramat, S., Handisurya, A., Kriehuber, E., Meneguzzi, G., Slupetzky, K., and Kirnbauer, R. (2003). Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. J Virol *77*, 13125-13135.

Shaheduzzaman, S., Krishnan, V., Petrovic, A., Bittner, M., Meltzer, P., Trent, J., Venkatesan, S., and Zeichner, S. (2002). Effects of HIV-1 Nef on cellular gene expression profiles. J Biomed Sci *9*, 82-96.

Shen, R., Fan, J. B., Campbell, D., Chang, W., Chen, J., Doucet, D., Yeakley, J., Bibikova, M., Wickham Garcia, E., McBride, C., *et al.* (2005). High-throughput SNP genotyping on universal bead arrays. Mutat Res *573*, 70-82.

Shiraki, K., Mukai, T., Okuno, T., Yamanishi, K., and Takahashi, M. (1991). Physicochemical characterization of human herpesvirus 6 infectivity. J Gen Virol *72 (Pt 1)*, 169-172.

Shyamsundar, R., Kim, Y. H., Higgins, J. P., Montgomery, K., Jorden, M., Sethuraman, A., van de Rijn, M., Botstein, D., Brown, P. O., and Pollack, J. R. (2005). A DNA microarray survey of gene expression in normal human tissues. Genome Biol *6*, R22.

Simmen, K. A., Singh, J., Luukkonen, B. G., Lopper, M., Bittner, A., Miller, N. E., Jackson, M. R., Compton, T., and Fruh, K. (2001). Global modulation of cellular transcription by human cytomegalovirus is initiated by viral glycoprotein B. Proc Natl Acad Sci U S A *98*, 7140-7145.

Simmons, A., Aluvihare, V., and McMichael, A. (2001). Nef triggers a transcriptional program in T cells imitating single-signal T cell activation and inducing HIV virulence mediators. Immunity *14*, 763-777.

Skrincosky, D., Hocknell, P., Whetter, L., Secchiero, P., Chandran, B., and Dewhurst, S. (2000). Identification and analysis of a novel heparin-binding glycoprotein encoded by human herpesvirus 7. J Virol *74*, 4530-4540.

Song, J. H., Kim, J. M., Kim, S. H., Kim, H. J., Lee, J. J., Sung, M. H., Hwang, S. Y., and Kim, T. S. (2003). Comparison of the gene expression profiles of monocytic versus granulocytic lineages of HL-60 leukemia cell differentiation by DNA microarray analysis. Life Sci *73*, 1705-1719.

Song, Y. J., and Stinski, M. F. (2002). Effect of the human cytomegalovirus IE86 protein on expression of E2F-responsive genes: a DNA microarray analysis. Proc Natl Acad Sci U S A *99*, 2836-2841.

Soukas, A., Cohen, P., Socci, N. D., and Friedman, J. M. (2000). Leptin-specific patterns of gene expression in white adipose tissue. Genes Dev *14*, 963-980.

Spear, P. G., and Longnecker, R. (2003). Herpesvirus entry: an update. J Virol *77*, 10179-10185.

Spector, D. H. (1996). Activation and regulation of human cytomegalovirus early genes. Intervirology *39*, 361-377.

Stamminger, T., Puchtler, E., and Fleckenstein, B. (1991). Discordant expression of the immediate-early 1 and 2 gene regions of human cytomegalovirus at early times after infection involves posttranscriptional processing events. J Virol *65*, 2273-2282.

Stasiak, P. C., and Mocarski, E. S. (1992). Transactivation of the cytomegalovirus ICP36 gene promoter requires the alpha gene product TRS1 in addition to IE1 and IE2. J Virol *66*, 1050-1058.

Steenman, M., Lamirault, G., Le Meur, N., Le Cunff, M., Escande, D., and Leger, J. J. (2005). Distinct molecular portraits of human failing hearts identified by dedicated cDNA microarrays. Eur J Heart Fail *7*, 157-165.

Steller, S., Angenendt, P., Cahill, D. J., Heuberger, S., Lehrach, H., and Kreutzberger, J. (2005). Bacterial protein microarrays for identification of new potential diagnostic markers for Neisseria meningitidis infections. Proteomics *5*, 2048-2055.

Stenberg, R. M. (1996). The human cytomegalovirus major immediate-early gene. Intervirology *39*, 343-349.

Stenberg, R. M., Depto, A. S., Fortney, J., and Nelson, J. A. (1989). Regulated expression of early and late RNAs and proteins from the human cytomegalovirus immediate-early gene region. J Virol *63*, 2699-2708.

Stingley, S. W., Ramirez, J. J., Aguilar, S. A., Simmen, K., Sandri-Goldin, R. M., Ghazal, P., and Wagner, E. K. (2000). Global analysis of herpes simplex virus type 1 transcription using an oligonucleotide-based DNA microarray. J Virol *74*, 9916-9927.

Stinski, M. F. (1978). Sequence of protein synthesis in cells infected by human cytomegalovirus: early and late virus-induced polypeptides. J Virol *26*, 686-701.

Stinski, M. F., Mocarski, E. S., and Thomsen, D. R. (1979). DNA of human cytomegalovirus: size heterogeneity and defectiveness resulting from serial undiluted passage. J Virol *31*, 231-239.

Stoughton, R. B. (2005). Applications of DNA microarrays in biology. Annu Rev Biochem *74*, 53-82.

Sullivan, D. C., Gray, W. L., Caughman, G. B., Robertson, A. T., and O'Callaghan, D. J. (1990). Temporal regulation of equine herpesvirus type 3 transcription. Virus Res *15*, 135-148.

Takeda, K., Okuno, T., Isegawa, Y., and Yamanishi, K. (1996). Identification of a variant A-specific neutralizing epitope on glycoprotein B (gB) of human herpesvirus-6 (HHV-6). Virology *222*, 176-183.

Talbot, S. G., Estilo, C., Maghami, E., Sarkaria, I. S., Pham, D. K., P, O. c., Socci, N. D., Ngai, I., Carlson, D., Ghossein, R., *et al.* (2005). Gene expression profiling allows distinction between primary and metastatic squamous cell carcinomas in the lung. Cancer Res *65*, 3063-3071.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E. S., and Golub, T. R. (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc Natl Acad Sci U S A *96*, 2907-2912.

Tan, F. L., Moravec, C. S., Li, J., Apperson-Hansen, C., McCarthy, P. M., Young, J. B., and Bond, M. (2002). The gene expression fingerprint of human heart failure. Proc Natl Acad Sci U S A *99*, 11387-11392.

Tanaka-Taya, K., Kondo, T., Nakagawa, N., Inagi, R., Miyoshi, H., Sunagawa, T., Okada, S., and Yamanishi, K. (2000). Reactivation of human herpesvirus 6 by infection of human herpesvirus 7. J Med Virol *60*, 284-289.

Tang, B. S., Chan, K. H., Cheng, V. C., Woo, P. C., Lau, S. K., Lam, C. C., Chan, T. L., Wu, A. K., Hung, I. F., Leung, S. Y., and Yuen, K. Y. (2005). Comparative host gene transcription by microarray analysis early after infection of the Huh7 cell line by severe acute respiratory syndrome coronavirus and human coronavirus 229E. J Virol *79*, 6180-6193.

Tavazoie, S., Hughes, J. D., Campbell, M. J., Cho, R. J., and Church, G. M. (1999). Systematic determination of genetic network architecture. Nat Genet *22*, 281-285.

Taylor, L. A., Carthy, C. M., Yang, D., Saad, K., Wong, D., Schreiner, G., Stanton, L. W., and McManus, B. M. (2000). Host gene regulation during coxsackievirus B3 infection in mice: assessment by microarrays. Circ Res *87*, 328-334.

Tenser, R. B., and Dunstan, M. E. (1980). Mechanisms of herpes simplex virus infectivity enhanced by ultracentrifugal inoculation. Infect Immun *30*, 193-197.

Tian, B., Zhang, Y., Luxon, B. A., Garofalo, R. P., Casola, A., Sinha, M., and Brasier, A. R. (2002). Identification of NF-kappaB-dependent gene networks in respiratory syncytial virus-infected cells. J Virol *76*, 6800-6814.

Torigoe, S., Kumamoto, T., Koide, W., Taya, K., and Yamanishi, K. (1995). Clinical manifestations associated with human herpesvirus 7 infection. Arch Dis Child *72*, 518-519.

Troesch, A., Nguyen, H., Miyada, C. G., Desvarenne, S., Gingeras, T. R., Kaplan, P. M., Cros, P., and Mabilat, C. (1999). Mycobacterium species identification and rifampin resistance testing with high-density DNA probe arrays. J Clin Microbiol *37*, 49-55.

Tsai, C. A., Lee, T. C., Ho, I. C., Yang, U. C., Chen, C. H., and Chen, J. J. (2005). Multi-class clustering and prediction in the analysis of microarray data. Math Biosci *193*, 79-100.

Tsavachidou, D., Podrzucki, W., Seykora, J., and Berger, S. L. (2001). Gene array analysis reveals changes in peripheral nervous system gene expression following stimuli that result in reactivation of latent herpes simplex virus type 1: induction of transcription factor Bcl-3. J Virol *75*, 9909-9917.

Tsunedomi, R., Iizuka, N., Hamamoto, Y., Uchimura, S., Miyamoto, T., Tamesa, T., Okada, T., Takemoto, N., Takashima, M., Sakamoto, K., *et al.* (2005). Patterns of expression of cytochrome P450 genes in progression of hepatitis C virus-associated hepatocellular carcinoma. Int J Oncol *27*, 661-667.

Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.

Ueda, K., Kusuhara, K., Okada, K., Miyazaki, C., Hidaka, Y., Tokugawa, K., and Yamanishi, K. (1994). Primary human herpesvirus 7 infection and exanthema subitum. Pediatr Infect Dis J *13*, 167-168.

Vahey, M. T., Nau, M. E., Jagodzinski, L. L., Yalley-Ogunro, J., Taubman, M., Michael, N. L., and Lewis, M. G. (2002). Impact of viral infection on the gene expression profiles of proliferating normal human peripheral blood mononuclear cells infected with HIV type 1 RF. AIDS Res Hum Retroviruses *18*, 179-192.

Van Gelder, R. N., von Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D., and Eberwine, J. H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A *87*, 1663-1667.

van 't Wout, A. B., Lehrman, G. K., Mikheeva, S. A., O'Keeffe, G. C., Katze, M. G., Bumgarner, R. E., Geiss, G. K., and Mullins, J. I. (2003). Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4(+)-T-cell lines. J Virol *77*, 1392-1402.

Vazquez-Ortiz, G., Pina-Sanchez, P., Hidalgo, A., Lazos, M., Moreno, J., Alvarado, I., Cruz, F., Hernandez, D. M., Perez-Plascencia, C., and Salcedo, M. (2005). [Global expression analysis in uterine cervical cancer: metabolic pathways and altered genes]. Rev Invest Clin *57*, 434-441.

Vertegaal, A. C., Kuiperij, H. B., van Laar, T., Scharnhorst, V., van der Eb, A. J., and Zantema, A. (2000). cDNA micro array identification of a gene differentially expressed in adenovirus type 5- versus type 12-transformed cells. FEBS Lett *487*, 151-155.

Wagner, A., Doerks, A., Aboud, M., Alonso, A., Tokino, T., Flugel, R. M., and Lochelt, M. (2000). Induction of cellular genes is mediated by the Bel1 transactivator in foamy virus-infected human cells. J Virol *74*, 4441-4447.

Wagner, E. K., Swanstrom, R. I., and Stafford, M. G. (1972). Transcription of the herpes simplex virus genome in human cells. J Virol *10*, 675-682.

Walduck, A., Rudel, T., and Meyer, T. F. (2004). Proteomic and gene profiling approaches to study host responses to bacterial infection. Curr Opin Microbiol *7*, 33-38.

Wang, D., Coscoy, L., Zylberberg, M., Avila, P. C., Boushey, H. A., Ganem, D., and DeRisi, J. L. (2002a). Microarray-based detection and genotyping of viral pathogens. Proc Natl Acad Sci U S A *99*, 15687-15692.

Wang, H., Hao, L., Shung, C. Y., Sunter, G., and Bisaro, D. M. (2003). Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. Plant Cell *15*, 3020-3032.

Wang, J., Delabie, J., Aasheim, H., Smeland, E., and Myklebost, O. (2002b). Clustering of the SOM easily reveals distinct gene expression patterns: results of a reanalysis of lymphoma study. BMC Bioinformatics *3*, 36.

Wang, Z. W., Sarmento, L., Wang, Y., Li, X. Q., Dhingra, V., Tseggai, T., Jiang, B., and Fu, Z. F. (2005). Attenuated rabies virus activates, while pathogenic rabies virus evades, the host innate immune responses in the central nervous system. J Virol *79*, 12554-12565.

Wathen, M. W., and Stinski, M. F. (1982). Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. J Virol *41*, 462-477.

Wenli, M., Yan, W., Hongmin, W., and Wenling, Z. (2004). An oligonucleotide microarray for the detection of vaccinia virus. Br J Biomed Sci *61*, 142-145.

Wilson, C. L., Pepper, S. D., Hey, Y., and Miller, C. J. (2004). Amplification protocols introduce systematic but reproducible errors into gene expression studies. Biotechniques *36*, 498-506.

Wilson, M., DeRisi, J., Kristensen, H. H., Imboden, P., Rane, S., Brown, P. O., and Schoolnik, G. K. (1999). Exploring drug-induced alterations in gene expression in Mycobacterium tuberculosis by microarray hybridization. Proc Natl Acad Sci U S A *96*, 12833-12838.

Wyatt, L. S., and Frenkel, N. (1992). Human herpesvirus 7 is a constitutive inhabitant of adult human saliva. J Virol *66*, 3206-3209.

Wyatt, L. S., Rodriguez, W. J., Balachandran, N., and Frenkel, N. (1991). Human herpesvirus 7: antigenic properties and prevalence in children and adults. J Virol *65*, 6260-6265.

Yadav, M., Nambiar, S., Khoo, S. P., and Yaacob, H. B. (1997). Detection of human herpesvirus 7 in salivary glands. Arch Oral Biol *42*, 559-567.

Yamanishi, K., Okuno, T., Shiraki, K., Takahashi, M., Kondo, T., Asano, Y., and Kurata, T. (1988). Identification of human herpesvirus-6 as a causal agent for exanthem subitum. Lancet *1*, 1065-1067.

Yang, I. V., Chen, E., Hasseman, J. P., Liang, W., Frank, B. C., Wang, S., Sharov, V., Saeed, A. I., White, J., Li, J., *et al.* (2002a). Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome Biol *3*, research0062.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002b). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res *30*, e15.

Yang, Y. H., and Speed, T. (2002). Design issues for cDNA microarray experiments. Nat Rev Genet *3*, 579-588.

Yasukawa, M., Hasegawa, A., Sakai, I., Ohminami, H., Arai, J., Kaneko, S., Yakushijin, Y., Maeyama, K., Nakashima, H., Arakaki, R., and Fujita, S. (1999). Down-regulation of CXCR4 by human herpesvirus 6 (HHV-6) and HHV-7. J Immunol *162*, 5417-5422.

Yasukawa, M., Inoue, Y., Ohminami, H., Sada, E., Miyake, K., Tohyama, T., Shimada, T., and Fujita, S. (1997). Human herpesvirus 7 infection of lymphoid and myeloid cell lines transduced with an adenovirus vector containing the CD4 gene. J Virol *71*, 1708-1712.

Yasukawa, M., Yakushijin, Y., Furukawa, M., and Fujita, S. (1993). Specificity analysis of human CD4+ T-cell clones directed against human herpesvirus 6 (HHV-6), HHV-7, and human cytomegalovirus. J Virol *67*, 6259-6264.

Yoshikawa, T., Asano, Y., Kobayashi, I., Nakashima, T., Yazaki, T., Suga, S., Ozaki, T., Wyatt, L. S., and Frenkel, N. (1993). Seroepidemiology of human herpesvirus 7 in healthy children and adults in Japan. J Med Virol *41*, 319-323.

Yuen, T., Wurmbach, E., Pfeffer, R. L., Ebersole, B. J., and Sealfon, S. C. (2002). Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. Nucleic Acids Res *30*, e48.

Zammatteo, N., Hamels, S., De Longueville, F., Alexandre, I., Gala, J. L., Brasseur, F., and Remacle, J. (2002). New chips for molecular biology and diagnostics. Biotechnol Annu Rev *8*, 85-101.

Zhang, Y., de Bolle, L., Aquaro, S., van Lommel, A., De Clercq, E., and Schols, D. (2001a). Productive infection of primary macrophages with human herpesvirus 7. J Virol *75*, 10511-10514.

Zhang, Y., Hatse, S., De Clercq, E., and Schols, D. (2000). CXC-chemokine receptor 4 is not a coreceptor for human herpesvirus 7 entry into CD4(+) T cells. J Virol *74*, 2011-2016.

Zhang, Y., Luxon, B. A., Casola, A., Garofalo, R. P., Jamaluddin, M., and Brasier, A. R. (2001b). Expression of respiratory syncytial virus-induced chemokine gene networks in lower airway epithelial cells revealed by cDNA microarrays. J Virol *75*, 9044-9058.

Zhao, H., Hastie, T., Whitfield, M. L., Borresen-Dale, A. L., and Jeffrey, S. S. (2002). Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. BMC Genomics *3*, 31.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., *et al.* (2001). Global analysis of protein activities using proteome chips. Science *293*, 2101-2105.

Zhu, H., Cong, J. P., Mamtora, G., Gingeras, T., and Shenk, T. (1998). Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. Proc Natl Acad Sci U S A *95*, 14470-14475.

Zhu, Z., Gershon, M. D., Ambron, R., Gabel, C., and Gershon, A. A. (1995). Infection of cells by varicella zoster virus: inhibition of viral entry by mannose 6-phosphate and heparin. Proc Natl Acad Sci U S A *92*, 3546-3550.